MODELING NON-FICKIAN TRANSPORT IN GEOLOGICAL FORMATIONS AS A CONTINUOUS TIME RANDOM WALK

Brian Berkowitz,[1] Andrea Cortis,[2] Marco Dentz,[3] and Harvey Scher[1]

Received 18 July 2005; revised 10 January 2006; accepted 2 May 2006; published 30 June 2006.

[1] Non-Fickian (or anomalous) transport of contaminants has been observed at field and laboratory scales in a wide variety of porous and fractured geological formations. Over many years a basic challenge to the hydrology community has been to develop a theoretical framework that quantitatively accounts for this widespread phenomenon. Recently, continuous time random walk (CTRW) formulations have been demonstrated to provide general and effective means to quantify non-Fickian transport. We introduce and develop the CTRW framework from its conceptual picture of transport through its mathematical development to applications relevant to laboratory- and field-scale systems. The CTRW approach contrasts with ones used extensively on the basis of the advection-dispersion equation and use of upscaling, volume averaging, and homogenization. We examine the underlying assumptions, scope, and differences of these approaches, as well as stochastic formulations, relative to CTRW. We argue why these methods have not been successful in fitting actual measurements. The CTRW has now been developed within the framework of partial differential equations and has been generalized to apply to nonstationary domains and interactions with immobile states (matrix effects). We survey models based on multirate mass transfer (mobile-immobile) and fractional derivatives and show their connection as subsets within the CTRW framework.

Citation: Berkowitz, B., A. Cortis, M. Dentz, and H. Scher (2006), Modeling non-Fickian transport in geological formations as a continuous time random walk, *Rev. Geophys.*, *44*, RG2003, doi:10.1029/2005RG000178.

1. INTRODUCTION

[2] The question of how to quantify contaminant transport in porous soils and rock has been the focus of research over several decades in hydrology, as well as in the closely related disciplines of soil science and petroleum engineering. Landmark tracer breakthrough experiments, dating mostly from the 1950s and 1960s, formed the basis for theoretical developments and analyses that considered almost exclusively the classical advection-dispersion equation (ADE) (see (74)). The ADE, and variants thereof, have continued to be used to this day as the principal means for considering and quantifying tracer transport in porous media.

[3] However, just how suitable is the ADE framework for describing tracer transport in natural porous media? Even early pioneering experiments reported the occurrence of systematic errors in fitting breakthrough curves (BTCs) using the classical ADE. In a series of careful and well-documented column experiments, *Scheidegger* [1959] observed that deviations in fits of the ADE to the BTCs could not be explained simply by the usual variability (error) in experimental measurements. *Scheidegger* [1959, p. 103] stated: ''The deviations are systematic which appears to point toward an additional, hitherto unknown, effect.'' *Aronofsky and Heller* [1957] also analyzed published tracer experiments and reported that systematic deviations arise between measurements and predictions using the ADE. Indeed, over the last 4 decades several other studies have pointed out specific and serious inadequacies in the applicability of the classical ADE, even with small-scale laboratory experiments on homogeneous samples. *Silliman and Simpson* [1987] demonstrated convincingly in laboratory experiments the scale dependency of the dispersivity coefficient; this is in stark contrast to the fundamental assumption that the dispersivity is a constant derived from the microgeometry of the porous medium. Such scale-dependent behavior is typical of ''non-Fickian'' or ''anomalous'' transport. These issues are by no means limited to laboratory-scale experiments; as we discuss in section 3,

[1]Department of Environmental Sciences and Energy Research, Weizmann Institute of Science, Rehovot, Israel.
[2]Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA.
[3]Department of Geotechnical Engineering and Geosciences, Technical University of Catalonia, Barcelona, Spain.

8755-1209/06/2005RG000178$15.00
Paper number 2005RG000178

deviations from the classical ADE behavior are even more significant in natural systems.

[4] While the ADE can treat ''homogeneous'' porous media under some conditions, such homogeneity rarely, if ever, exists. The heterogeneity of natural geological formations at a wide range of scales necessitates consideration of more sophisticated transport theories. The purpose of this review is to introduce an effective theoretical and practical basis that accounts for the behavior of solute transport in heterogeneous media in field, laboratory, and numerical experiments based on the continuous time random walk (CTRW) formalism. Here we shall discuss all aspects of CTRW, focusing on its relevance and applicability as a framework that has dealt successfully with the phenomenon of anomalous transport. We then discuss and critique, relative to CTRW, other approaches found in the literature. We show that some of the latter are specialized cases within the CTRW framework. Throughout this review we use the terms ''tracer,'' ''particle,'' ''contaminant,'' and ''solute'' interchangeably.

[5] Why is the conceptual picture underlying the classical ADE formulation limited? What is missing from this picture describing the transport behavior of a contaminant in a natural porous medium? The answers to these questions lie in the basic recognition that in all geological formations, heterogeneities are present at all scales, from the submillimeter pore scale to the basin scale itself. We emphasize that the term ''heterogeneities'' can refer to variations in the distribution of the geometrical properties (e.g., porosity and hydraulic conductivity), as well as to variations in the biogeochemical properties of the medium, all of which can affect tracer transport. Three points of conceptual understanding can be drawn immediately from this fact: (1) The high degree of variability in these heterogeneities rules out, a priori, the possibility of obtaining complete knowledge of the pore space in which fluids and contaminants are transported. (2) The paths traveled by a contaminant in an aquifer are strongly influenced by the heterogeneities of the geological formation, as well as by the initial and boundary conditions (BCs), which determine the underlying flow field. (3) Tracer migration is sensitive to heterogeneities at all scales, so that we should not be surprised that small-scale heterogeneities can significantly affect large-scale behavior.

[6] The key consequence of these points is a critical consideration of the idea of ''homogeneity'' of the medium for the purposes of modeling transport and/or defining ''effective'' transport parameters. It has been shown that even carefully packed, laboratory-scale flow cells and columns containing porous media contain ''heterogeneities.'' Studies using magnetic resonance imaging to visualize flow conditions within ''homogeneous'' geological materials in laboratory-scale, column experiments report the existence of preferential flow paths, which strongly influence both water flow and tracer transport [e.g., *Hoffman et al.*, 1996; *Oswald et al.*, 1997]. These paths occur because of the presence of macrostructures (caused, e.g., from bridging effects) as well as by microstructures that reflect grain-size heterogeneities.

[7] At the field scale the issue of ''homogenization'' of course arises. Do the above mentioned laboratory-scale heterogeneities simply average out and become insignificant at large scales? If so, how large is ''large''? As already noted above, heterogeneities are present at all scales. Thus, for example, the existence of preferential flow paths has been reported even in apparently ''structureless'' soils at the field scale [e.g., *Ghodrati and Jury*, 1992].

[8] The nature of contaminant transport in geological materials is thus linked inextricably to the extent and scale dependence of heterogeneities. The transport that is anomalous, or non-Fickian, occurs when the contaminant encounters, at each scale, a sufficiently broad spectrum of velocities and stagnant areas resulting from the heterogeneities. In addition to the strong influence of preferential flow paths and slow flow or diffusion-dominated regions, tracers can in many cases be affected by biogeochemical heterogeneities; these heterogeneities enable a wide range of reaction processes that (temporarily or permanently) delay the advance of a tracer. Non-Fickian behavior is fundamentally different from Fickian transport, which is the usual assumption invoked, explicitly or implicitly, for application of the classical ADE and many of its variants. These other treatments, and many stochastic approaches, focus on definition of an effective ''macrodispersion'' parameter at any given scale of interest. This issue is discussed in detail in section 6.

[9] The concept of anomalous transport was first introduced by *Montroll and Scher* [1973] and *Scher and Montroll* [1975] and has subsequently been shown to have ubiquitous applicability to transport and diffusion in disordered systems. The original application of the concept and its quantitative predictions [*Scher and Montroll*, 1975], discussed in detail in section 2.5, was in the field of electronic transport in amorphous semiconductors [*Tiedje*, 1984] and in polymeric media [*Bos and Burland*, 1987] where it was well confirmed by an extensive number of experimental studies. Subsequent use of these ideas expanded into many areas, e.g., the anomalous diffusion of defects as a basis for the universal stretched exponential relaxation behavior in a diverse number of materials [*Shlesinger*, 1988]. The present application to contaminant transport in geological formations is a uniquely rich example, as the full chemical plume and BTCs can be measured directly; this application has represented a new level of confirmation and further development of the theory.

[10] Here and throughout, we use the terms ''anomalous'' and ''non-Fickian'' interchangeably to denote any transport behavior which differs from that described by the classical ADE. (Some authors refer to this behavior as ''preasymptotic'' or ''preergodic.'') The ADE describes Fickian behavior in the sense that mechanical dispersion is assumed to be quantifiable by a macroscopic form of Fick's law, and the resulting temporal and spatial concentration distributions of an initial pulse are equivalent to a normal, otherwise known as a Gaussian, distribution.

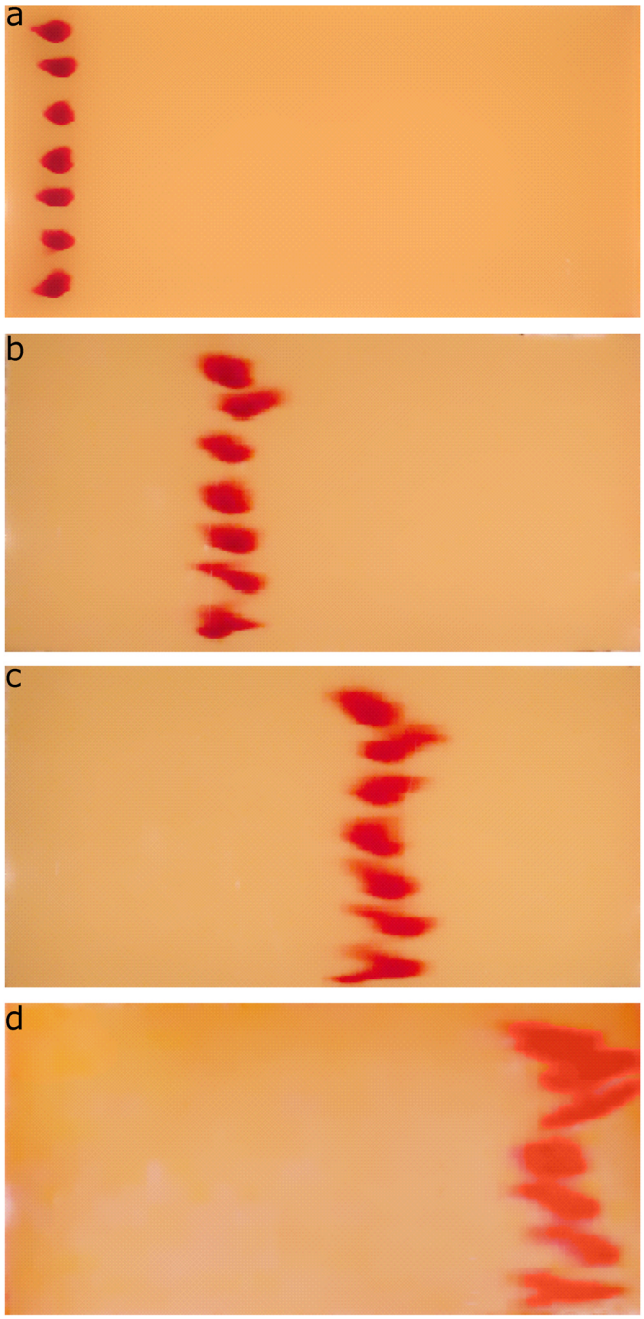

Figure 1. Photographs of a homogeneous, saturated sand pack with seven dye tracer point injections being transported, under constant flow of 53 mL/min, from left to right. Times at (a) t = 20, (b) t = 105, (c) t = 172, and (d) t = 255 min after injection. Internal dimensions of the flow cell are 86 cm (length), 45 cm (height), and 10 cm (width). Reprinted from *Levy and Berkowitz* [2003]. Copyright 2003, with permission from Elsevier.

[11] To illustrate just how frequently non-Fickian transport occurs, consider tracer migration through an "homogeneous," fully saturated sand pack. As indicated above, preferential paths for fluid flow and tracer transport are present even in these conditions, as shown, for example, in Figure 1. Contrary to Fickian transport the individual dye plumes are not symmetrical ellipses, nor are the different plumes identical to each other. Moreover, measurements of tracer BTCs in such "homogeneous," meter-length flow cells have been shown to display "anomalous" early time arrivals (i.e., later than Fickian) and late time tails [*Levy and Berkowitz*, 2003]. Detailed analysis (see section 3.4.2) shows that the motion and spreading of these chemical plumes are characterized by distinct temporal scaling; that is, the time dependence of the spatial moments does not correspond to a normal (or Gaussian) distribution. In sections 2–5 we discuss the CTRW approach to these transport phenomena, and in sections 6 and 7 we contrast it with other approaches that have been fully discussed in the literature.

2. CONTINUOUS TIME RANDOM WALK FRAMEWORK

[12] Two well-studied, generic geological media that possess heterogeneities on a very wide range of spatial scales are porous sedimentary rock and "random" fracture networks (RFN) in low-permeability rock. At the field scale a reasonable definition of the macroscopic characteristics (e.g., individual facies) of these geological formations can be feasible, thus enabling, at a sufficiently coarse resolution, modeling of flow and transport conditioned on these features. However, in practice, there is always some scale, ℓ, below which ($y < \ell$) heterogeneities are unresolved. The omnipresent question is, Can one justify the use of average local properties (e.g., mean velocity and dispersion) at the scale ℓ, or does the range of unresolved heterogeneities $y < \ell$ have a key influence on overall transport behavior? The answer is very often a practical one, not an intrinsic one, depending on the width of the distribution of material properties for $y < \ell$.

[13] To account for the effect of a sufficiently broad (statistical) distribution of material properties (e.g., of permeabilities) on the overall transport, one must consider a probabilistic approach that will generate a probability density function (pdf) describing key features of the transport. This pdf, denoted $\psi(\mathbf{s}, t)$, is discussed thoroughly starting in section 2.3. The effects of multiscale heterogeneities on contaminant transport patterns are significant, and consideration only of the mean transport behavior, e.g., the spatial moments of the concentration distribution, is not sufficient. An essential input to the calculation of the field-scale transport is the plume motion and/or BTC across the ℓ scale. The CTRW is a probabilistic approach for calculating the latter based on a pdf of transition times (see section 2.1) generated by the range of heterogeneities. The nature of the transport, non-Fickian or Fickian, is determined by the functional shape of the pdf.

2.1. Conceptual Picture: Tracer Transitions

[14] Contaminant motion in geological formations can be treated by considering particles (which represent, e.g., dissolved solutes) undergoing various types of transitions. These transitions encompass both the displacement due to structure and heterogeneity as well as the time taken to

make the particle movement between, e.g., pores or fracture intersections. We conceptualize transport as a series of such particle transitions with a focus on retaining the full distribution of the transition times. The variability in the hydraulic and geochemical properties of the geological domain cause a variety of particle transitions at velocity changes within the flow field, between flowing and stagnant zones, between mobile and immobile states, between macropores and micropores, between fractures and adjacent host rock, and by changes in advective paths at fracture and macropore intersections. This picture of motion by transitions will be referred to as "CTRW theory," whereas the mathematical formalism used to implement the motion will be called the "CTRW framework."

[15] Each transition can be quantified as $w(\mathbf{s}, \mathbf{s}')$, the rate of particle transfer to position $\mathbf{s}$ from $\mathbf{s}'$, and can be considered on any spatial scale, e.g., on a pore scale between pore positions through an interpore throat (i.e., a "tube"). A multiple rate approach considers the range of these rates $\{w\}$. At this point one can see the basic problem in working with average rates in some representative region or volume, traditionally referred to as a "representative elementary volume" (REV). A particle "encounter" with a sparsely distributed, very small rate w_0 can have a large impact on the overall transport, but w_0 can be entirely absent in a REV average of $\{w\}$; other problems with the REV have been pointed out previously in the literature. Thus the details of the distribution of $\{w\}$, or as we will show the ensemble average of the $\{w\}$ over all the configurations of a specific system, are key to the nature of the transport. The variation in spatial displacements in the distribution of $w(\mathbf{s}, \mathbf{s}')$ is small in the type of transport typically encountered in geological formations. For example, the "tube" lengths in the pore-scale model above have a narrow distribution. However, the variation in rates (i.e., values of $w(\mathbf{s}, \mathbf{s}')$), which is governed by the velocity spectrum of the flow field, is very large for highly disordered media; for example, the fluid flow distribution in the "tubes" governs the transit time between pore sites (see section 3.8). Hence the temporal distribution of the pdf (i.e., the range of $[w(\mathbf{s}, \mathbf{s}')]^{-1}$) discussed at the beginning of section 2 dominates the nature of the transport. The emphasis on temporal aspects of particle transport, induced by the spatial heterogeneity, is a key feature of the CTRW approach. We consider the significance of this emphasis further when we contrast CTRW to the usual ADE framework in section 6.1.

[16] For now, we picture tracer transport as a series of discrete (in space) transitions. These can be defined naturally, for example, as transitions through a fracture network between fracture intersections. If we carefully retain the temporal distribution of these transitions, this picture can be expanded easily to a continuous-in-space formulation, as we shall show in section 2.6.

2.2. Basic Formulation of Transport

[17] Our point of departure is a general framework that can encompass all of the processes enumerated above as special cases and reduce to the ADE for a "perfectly homogeneous" medium. Defining this overarching framework is a transport equation incorporating the full range of $\{w\}$ for any given realization of the domain,

$$\frac{\partial C(\mathbf{s}, t)}{\partial t} = -\sum_{\mathbf{s}'} w(\mathbf{s}', \mathbf{s}) C(\mathbf{s}, t) + \sum_{\mathbf{s}'} w(\mathbf{s}, \mathbf{s}') C(\mathbf{s}', t), \quad (1)$$

where $C(\mathbf{s}, t)$ is the normalized particle concentration or probability at point $\mathbf{s}$ and time t in a specific realization of the domain and the dimension of $\Sigma_{\mathbf{s}} w$ is reciprocal time.

[18] Equation (1) expresses a conservation of mass at each site $\mathbf{s}$ and describes the rate of concentration change at $\mathbf{s}$ as a function of the distribution of probabilities of moving from $\mathbf{s}$ to $\mathbf{s}'$ and from $\mathbf{s}'$ to $\mathbf{s}$. This equation is known as the "master equation" (ME) [*Oppenheim et al.*, 1977; *Shlesinger*, 1996]. It has been utilized widely in the physics and chemistry literature, e.g., electron hopping in random systems [e.g., *Klafter and Silbey*, 1980a].

[19] In most of the applications considered here, the transition rates describe the effects of the velocity field on the particle motion. It is important to point out that the transport equation (1) does not separate the effects of the varying velocity field into an advective and dispersive part of the motion.

[20] Specification of $w(\mathbf{s}, \mathbf{s}')$ involves detailed knowledge of the system, i.e., characterization of the heterogeneities on all length scales that influence the calculation of the flow field. Below the ℓ scale we must resort to a statistical description of this subdomain and hence to a distribution of $\{w\}$. To realize this probabilistic approach, we consider the ensemble average of (1), which can be shown [*Klafter and Silbey*, 1980b] to be of the form

$$\frac{\partial c(\mathbf{s}, t)}{\partial t} = -\sum_{\mathbf{s}'} \int_0^t \phi(\mathbf{s}' - \mathbf{s}, t - t') c(\mathbf{s}, t') dt' + \sum_{\mathbf{s}'} \int_0^t \phi(\mathbf{s} - \mathbf{s}', t - t') c(\mathbf{s}', t') dt', \quad (2)$$

where $c(\mathbf{s}, t)$ is the mean, ensemble-averaged, normalized concentration and $\phi(\mathbf{s}, t)$ is defined in (7). The form of (2) is a "generalized master equation" (GME) which in contrast to (1) is nonlocal in time, that is, (2) contains an integral over time requiring knowledge of the past state of the concentration.

[21] The ensemble average of a set of local (in time) kinetic equations (e.g., equation (1)) for a disordered system leads to a nonlocal transport equation, because all of the $\{w\}$ are made available to each site and the role of $w(\mathbf{s}', \mathbf{s})$ is replaced by a function of time which depends on a distribution of transit times between sites. Hence the ensemble average of any set of equations describing the dynamics of a physical model of a disordered system will lead to a nonlocal equation. (For a simple example see section 7.1 (especially (90) and (91)).) The various nonlocal transport equations often have similar form, but there is no intrinsic relation between them as each depends on the physical

model that generated them (see section 6.2). In fact, the use of nonlocal equations for a broad class of transport problems has a long history [e.g., *Zwanzig*, 1960; *Mori*, 1965, and references therein]. Applications more specifically related to the present ones include those of *Kenkre et al.* [1973], *Montroll and Scher* [1973], *Scher and Lax* [1973a], *Shlesinger* [1974], *Scher and Montroll* [1975], and *Klafter and Silbey* [1980a].

[22] The transition rates in (2) are time-dependent but stationary, depending only on the difference $\mathbf{s} - \mathbf{s}'$. Hence, depending on available knowledge of the system, ℓ can range from meters to tens and hundreds of meters. As in (1), note that in (2) there is no separation between an advective and dispersive part of the motion.

2.3. CTRW Transport Equations

[23] Using the Laplace transform, it can be shown [*Kenkre et al.*, 1973; *Shlesinger*, 1974] that the GME (2) is completely equivalent to a CTRW (see Appendix A)

$$R(\mathbf{s}, t) = \sum_{\mathbf{s}'} \int_0^t \psi(\mathbf{s} - \mathbf{s}', t - t') R(\mathbf{s}', t') dt', \tag{3}$$

where $R(\mathbf{s}, t)$ is the probability per time for a walker to just arrive at site $\mathbf{s}$ at time t and $\psi(\mathbf{s}, t)$ is the probability per time for a displacement $\mathbf{s}$ with a difference of arrival times of t. The initial condition for $R(\mathbf{s}, t)$ is $\delta_{\mathbf{s},0}\,\delta(t - 0^+)$, which can be appended to (3). The $\psi(\mathbf{s}, t)$ is the basic pdf discussed at the beginning of section 2; $\psi(\mathbf{s}, t)$ determines the nature of the transport, as will be considered in applications below.

[24] A random walk with continuous time was introduced by *Montroll and Weiss* [1965] using a distribution $\psi(t)$ (see (5)) for the step time. The generalization of the formalism, i.e., the appearance of equation (3) with the joint distribution $\psi(\mathbf{s}, t)$ and labeled "CTRW," and the physical application to transport, was first given by *Scher and Lax* [1973a]. Equation (3) describes a semi-Markovian process, Markovian in space but not in time, which accounts for memory in particle transitions. The CTRW reduces to a Markovian random walk (see Appendix A) for a single rate, i.e., $\psi(\mathbf{s}, t) = w(\mathbf{s})\exp(-Wt)$ and $W \equiv \Sigma_{\mathbf{s}}\, w(\mathbf{s})$ (see section 2.5).

[25] The correspondence between (2) and (3) is

$$c(\mathbf{s}, t) = \int_0^t \Psi(t - t') R(\mathbf{s}, t') dt', \tag{4}$$

where

$$\Psi(t) = 1 - \int_0^t \psi(t') dt' \tag{5}$$

is the probability for a walker to remain on a site,

$$\psi(t) \equiv \sum_{\mathbf{s}} \psi(\mathbf{s}, t) \tag{6}$$

$$\tilde{\phi}(\mathbf{s}, u) = \frac{u\tilde{\psi}(\mathbf{s}, u)}{1 - \tilde{\psi}(u)}, \tag{7}$$

where the Laplace transform ($\mathcal{L}$) of a function $f(t)$ is denoted by $\tilde{f}(u)$.

[26] Equations (3)–(5) are in the form of a convolution in space and time and can therefore be solved using Fourier ($\mathcal{F}$) and Laplace transforms [*Scher and Lax*, 1973a]. The general solution, for periodic BCs in a lattice of size N (with site positions $\mathbf{s} = \sum_{j=1}^{3} s_j a_j$, $s_j = 1, 2, 3, \ldots, N$, and a_j = lattice constant), is

$$\mathcal{C}(\mathbf{k}, u) = \frac{1 - \tilde{\psi}(u)}{u} \frac{1}{1 - \Lambda(\mathbf{k}, u)} \tag{8}$$

where $\mathcal{C}(\mathbf{k}, u)$ and $\Lambda(\mathbf{k}, u)$ are the Fourier transforms of $\tilde{c}(\mathbf{s}, u)$ and $\tilde{\psi}(\mathbf{s}, u)$, respectively, where for each component of $\mathbf{k}$ the range of k values is $k_i = 2\pi l_i/N$, and l_i is an integer, $-(N-1)/2 \le l_i \le (N-1)/2$, for odd N. An input of $\psi(\mathbf{s}, t)$ in Λ in (8) leads to the determination of $c(\mathbf{s}, t)$, which represents the tracer plume concentration after a $\mathcal{F}$ and $\mathcal{L}$ inversion. Because N is considered to be very large, the solution (8) is regarded as being in the infinite domain, vanishing at infinity. Solutions for a bounded domain and for more general boundary conditions are developed in sections 3 and 4. Note also that as N is very large, the lattice constant can be arbitrarily small (we develop the continuum limit in section 2.6), and, for example, $\psi(\mathbf{s}, t)$ can have a range of many lattice sites. The lattice thus acts as a "scaffold" to determine the solution (8) and does not confine the spatial distribution of the plume.

[27] In addition to the determination of the concentration plume (8), another key function in CTRW is the first-passage time distribution $F(\mathbf{s}, t)$, the probability density for a walker starting at the origin to reach $\mathbf{s}$ for the first time. The solution in (8) is for periodic BCs; experiments and observations often call for an absorbing BC or exit plane. The main measurement for these experiments is the BTC, which is equivalent to the $F(\mathbf{s}, t)$ evaluated on a plane (e.g., $s_1 = L$).

[28] The implicit relation for $F(\mathbf{s}, t)$ is

$$R(\mathbf{s}, t) = \delta_{\mathbf{s},0}\delta(t - 0^+) + \int_0^t F(\mathbf{s}, t') R(0, t - t') dt', \tag{9}$$

which states that the walker arrives at $\mathbf{s}$ for the first time at time t' and in the remaining time $t - t'$, the walker can visit and leave $\mathbf{s}$ an arbitrary number of times but ends at $\mathbf{s}$. Equation (9) contains a convolution in time, and so in Laplace space the relation becomes an algebraic one that is solved easily:

$$\tilde{R}(\mathbf{s}, u) = \delta_{\mathbf{s},0} + \tilde{F}(\mathbf{s}, u)\tilde{R}(0, u); \tag{10}$$

thus

$$\tilde{F}(\mathbf{s}, u) = \frac{\tilde{R}(\mathbf{s}, u) - \delta_{\mathbf{s},0}}{\tilde{R}(0, u)}. \tag{11}$$

To obtain a BTC, we first consider that the walk starts from a plane, e.g., $s_1 = 0$, and evaluate $F(\mathbf{s}, t)$ at a fixed distance $s_1 = L$ by summing over the other s_i directions ($i > 1$),

$$f_B(t) \equiv \Sigma_{s_2, s_3} F(s_1 = L, s_2, s_3, t). \tag{12}$$

A BTC in this paper will generally refer to the cumulative of $f_B(t)$ over time.

2.4. Numerical Inversion of Laplace Transforms

[29] The analytical solutions developed in section 2.3, as well as in sections 3 and 4, are of limited use in their Laplace space form. We therefore need to invert these solutions to the time domain. The inversion for the $\mathcal{L}$ involves finding the solution $g(t)$ of an integral equation of the first kind [*Krylov and Skoblya*, 1977]:

$$\int_0^\infty g(t)e^{-ut}dt = G(u), \tag{13}$$

where $G(u)$ is a given function of the complex parameter u.

[30] In a few important cases we can develop $G(u)$ in an asymptotic form, for small u behavior, and solve for $g(t)$ for large t by analytic means (see section 3). In general, we must resort to numerical means. We use the *de Hoog et al.* [1982] Laplace inversion algorithm, which makes use of complex-valued Laplace parameters.

[31] The algorithm works as follows: Suppose that one is interested in the solution for a range of times spanning from $t_{\min}$ to $t_{\max}$. This range of times is then discretized in a time vector of arbitrary length N. It was observed that simultaneous inversion for times covering several orders of magnitudes gives inaccurate results for the small times. Therefore the algorithm splits the time vector for which we want to obtain the concentrations into sections of the same order of magnitude (usually a logarithmic cycle), and the individual sections are inverted at a given time. So if the time vector spans n orders of magnitude, the inversion process will run n times. It is often a good idea to discretize the time vector in such a way that its values are equally spaced on a logarithmic scale.

[32] The solution of the complex-valued partial differential equation (pde) (e.g., (24), (33), or (73)) must be known for a series of u values determined by the expression

$$u_{k,j} = \frac{1}{2T_k}\left(-\log_{10}\epsilon + j2\pi\sqrt{-1}\right), \tag{14}$$

where T_k is the maximum time for the kth piece of time vector ($k = 1, \ldots, n$), and $j = 1, \ldots, m$. The ϵ parameter is typically of the order of 10^{-9}, and m is the number of terms in the Fourier series expansion (obtained from the inversion integral using the trapezoidal rule), typically in the range of 20 to 40. Note that, for each of the n pieces, the complex-valued $u_{k,j}$ in the above formula have an invariant real part and a variable imaginary part where the Fourier series expansion is calculated (with the use of an accelerated convergence method of the algorithm [*de Hoog et al.*, 1982]). Practically, the user provides the full range of times and the numerical inversion subroutine will call the pde solver $n \times m$ times, one for each of the $u_{k,j}$ values determined by the above procedure. The output is a series of concentration values at the discretized times in the time vector.

[33] On the basis of the solutions discussed herein and the work of *Dentz et al.* [2004] and *Cortis et al.* [2004b], we have developed a CTRW "toolbox" [*Cortis and Berkowitz*, 2005] that provides a collection of easy-to-use MATLAB scripts and functions to calculate the full temporal and spatial behavior of a migrating tracer. The CTRW toolbox is freely available at http://www.weizmann.ac.il/ESER/People/Brian/CTRW.

2.5. Relation Between $\psi(\mathbf{s}, t)$ and $w(\mathbf{s}', \mathbf{s})$

[34] The identification of $\psi(\mathbf{s}, t)$ lies at the heart of the CTRW formulation, as it defines the nature of the plume properties. We will show in section 3 that the distribution of t in $\psi(\mathbf{s}, t)$ is the essential feature because it captures the often large range of the values of $w(\mathbf{s}', \mathbf{s})$. Clearly, the underlying permeability distribution of the domain gives rise to the velocity distribution, which thus includes naturally all correlations affecting transport. In the CTRW formulation the velocity distribution is used to define the spectrum of w and hence $\psi(\mathbf{s}, t)$. In section 2.1 we outlined the qualitative features that can give rise to anomalous transport: the sufficient encounter of a tracer with a relatively rare slow rate (w_0) or zone. We now quantify this qualitative statement by showing the relation between $\psi(\mathbf{s}, t)$ and $w(\mathbf{s}', \mathbf{s})$.

[35] To show the general relation between $\psi(\mathbf{s}, t)$ and $w(\mathbf{s}', \mathbf{s})$, we rewrite the master equation (1) (which is for a specific representation) as a random walk equation to obtain a transition length and time distribution $\psi_{\mathbf{s}',\mathbf{s}}(t)$ of the form

$$\psi_{\mathbf{s}',\mathbf{s}}(t) = w(\mathbf{s}', \mathbf{s}) \exp\left[-t\sum_{\mathbf{s}''} w(\mathbf{s}'', \mathbf{s})\right]. \tag{15}$$

In this equation (derived in Appendix B), $\psi_{\mathbf{s}',\mathbf{s}}(t)$ plays the role of $\psi(\mathbf{s}, t)$, but it is dependent on the location; that is, each neighborhood is different in a specific representation.

[36] Furthermore, we can then write

$$\psi_{\mathbf{s}}(t) = \sum_{\mathbf{s}'} w(\mathbf{s}', \mathbf{s}) \ \exp\left[-t\sum_{\mathbf{s}'} w(\mathbf{s}', \mathbf{s})\right] = -\frac{dQ_{\mathbf{s}}}{dt}, \tag{16}$$

where

$$Q_{\mathbf{s}} \equiv \exp\left[-t\sum_{\mathbf{s}'} w(\mathbf{s}', \mathbf{s})\right] \tag{17}$$

$$\psi(t) = -\frac{d}{dt}[[Q_{\mathbf{s}}]], \tag{18}$$

with double brackets denoting an ensemble average. In general, it is very difficult to carry out the ensemble average in (18), e.g., over a specified flow distribution. When the transitions are within a specific configuration of random spatial locations or sites, one can calculate the ensemble average exactly; see below for the results of this calculation and section 3.5 for a hydrological application.

[37] The simplest possible model for $\psi(t)$ corresponds to the use of an average rate or an ordered case of lattice site positions with $w(\mathbf{s}', \mathbf{s}) = w(\mathbf{s}' - \mathbf{s})$ and the total rate $W = \Sigma_{\mathbf{s}'} w(\mathbf{s}', \mathbf{s})$ in (16), i.e., an exponential of the form $\psi(t) = W\exp(-Wt)$ and $\tilde{\psi}(u) = W/(W + u)$. In this ordered or averaged case, (15) can be written as $\psi(\mathbf{s}, t) = w(\mathbf{s})\exp(-Wt)$, and therefore (7) becomes $\tilde{\phi}(\mathbf{s}, u) = w(\mathbf{s})$, independent of u.

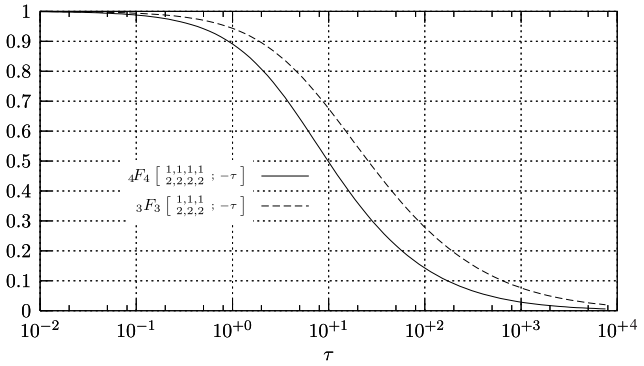

Figure 2. Plot of the ${}_3F_3$ and ${}_4F_4$ functions in (21) versus dimensionless time τ [after *Cortis et al.*, 2004b].

Hence (2) becomes local in t and is equal to (1) with rates $w(\mathbf{s}' - \mathbf{s})$; see also Appendix A.

[38] In contrast, a u dependence of $\tilde{\phi}(\mathbf{s}' - \mathbf{s}, u)$ derives from a range of rates, which characterize a disordered system. To generate this range, we can maintain a $w(\mathbf{s}' - \mathbf{s})$ but now have a random spatial distribution of sites. We consider the ensemble average over all possible distributions of sites in a domain with site density N_s of $[[Q_\mathbf{s}]]$ in (18), evaluating it analytically to obtain [*Scher and Lax*, 1973b]

$$[[Q_\mathbf{s}]] = \exp\left\{-N_s \int [1 - \exp(-w(\mathbf{s})t)]d^3s\right\}. \tag{19}$$

[39] To evaluate (19) further, we choose a standard transition rate between local sites,

$$w(\mathbf{s}' - \mathbf{s}) = w_m \exp(-|\mathbf{s}' - \mathbf{s}|/r_o), \tag{20}$$

where r_o is the range of the transitions and w_m is the maximum rate. Substituting (20) into (19) and then (18), we derive [*Cortis et al.*, 2004a],

$$\frac{\psi(\tau)}{w_m} = \zeta_3 F_3\left[\begin{matrix}1,1,1\\2,2,2\end{matrix};-\tau\right] e^{-\zeta\tau_4 F_4\left[\begin{smallmatrix}1,1,1,1\\2,2,2,2\end{smallmatrix};-\tau\right]}, \tag{21}$$

where $\tau \equiv w_m t$, $\zeta \equiv 4\pi N_s r_o^3$, and ${}_pF_q$ is the generalized hypergeometric function in standard notation as defined by *Abramowitz and Stegun* [1970] and shown in Figure 2. Equation (21) will be referred to as the η function, as discussed in section 3.5. Figure 3 illustrates the behavior of (21); one can observe that for the lower values of ζ, $\psi(\tau)$ varies slowly, i.e., as a power law $\sim\tau^{-1-\beta}$, where β is a parameter that is discussed at length below. There are large ranges of τ where β can have a value between 0 and 2; this will be an important aspect of the transport model (see section 3). The parameter ζ can be seen to be a measure of the disorder in the system or, more precisely, the range of rates.

[40] For transition rates of the form (20) the range of rates depends on an interplay between the spatial extent of the transition, r_o, and the average site separation, r_N, where $4\pi N_s/3 \equiv r_N^{-3}$ [*Scher and Lax*, 1973b]. The parameter ζ can be written as $\zeta = 3(r_o/r_N)^3$, and for $r_o \ll r_N$ the rates are sensitive to the separations between sites and hence generate a wide range of rates. For each value of ζ, there is a τ_c such that the effective β is greater than 2 for $\tau > \tau_c$, and the transport evolves to normal Gaussian behavior. Thus β should be seen as an effective parameter describing the asymptotic behavior of $\psi(t)$ over a time range corresponding to the duration of the observation. The large τ or asymptotic behavior of (21) is [*Scher and Lax*, 1973b]

$$\frac{\psi(\tau)}{w_m} = \zeta\left[(\ln c\tau)^2 + a_2\right] \cdot \frac{\exp\left\{-\frac{1}{3}\zeta\left[(\ln c\tau)^3 + 3a_2 \ln c\tau + 2a_3\right]\right\}}{\tau}, \tag{22}$$

where c, a_2, and a_3 are constants.

[41] In the CTRW formalism the term asymptotic refers to the large τ limit (which effectively can set in for values of $\tau \gtrsim 10$) or in Laplace space the small μ ($\equiv u/w_m$) limit, e.g., $\mu < 0.1$. In practice therefore the asymptotic domain covers most laboratory and field measurements. Note that in the literature that deals with dispersion (see section 3.7) the term "asymptotic" typically refers to the regime where an effective macrodispersion can be defined; to make a clear distinction, we will henceforth refer to this latter regime as the "macrodispersion regime."

[42] At large enough time the transport in a real physical system will become normal if the medium has a largest heterogeneity scale; in other words, at length scales sufficiently greater than the largest heterogeneity scale, the medium acts as a more "homogeneous" one. In terms of CTRW the effective β has a slowly varying time dependence (on $\ln\tau$ [see *Scher et al.*, 2002a]) and increases as $\tau \to \infty$. We shall document this behavior in detail with a simpler form of $\psi(\tau)$, the truncated power law function (61) in section 3.7. In section 3.5 we shall develop a hydrological application of (21), with a coarse graining of

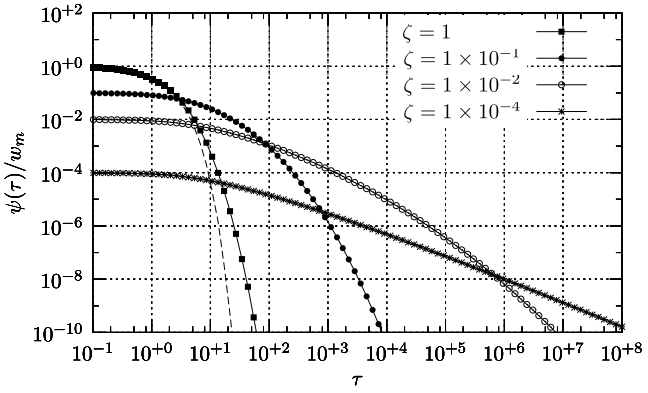

Figure 3. Evolution of the dimensionless transition probability $\psi(\tau)/w_m$ versus dimensionless time τ in (21) for different values of the ζ parameter. The dashed line represents the exponential limit of the function for $\zeta = 1.5$. Modified and reprinted with permission from *Cortis et al.* [2004a] (http://dx.doi.org/10.1103/PhysRevE.70.041108). Copyright 2004 by the American Physical Society.

a porous medium and an account of local velocity fluctuations [*Cortis et al.*, 2004a].

2.6. CTRW in a Partial Differential Equation Form

[43] In section 2.1 we indicated that the CTRW formalism can be extended easily to continuous-in-space formulations. We now discuss a convenient pde form of the CTRW equations.

[44] Many transport problems that are modeled with CTRW have a $\psi(\mathbf{s}, t)$ that allows all spatial moments to exist, while in many cases not even the first temporal moment exists (or it is very large). We discuss this issue in section 3.3. As a consequence a good approximation to $c(\mathbf{s}', t)$ is a Taylor expansion over the finite range of the transition rates:

$$c(\mathbf{s}', t) \approx c(\mathbf{s}, t) + (\mathbf{s}' - \mathbf{s}) \cdot \nabla c(\mathbf{s}, t) + \frac{1}{2}(\mathbf{s}' - \mathbf{s})(\mathbf{s}' - \mathbf{s}) : \nabla\nabla c(\mathbf{s}, t), \quad (23)$$

with the dyadic symbol (colon) denoting a tensor product (see Appendix C for an example) [*Berkowitz et al.*, 2002].

[45] It can be shown that insertion of (23) into (2) yields in Laplace space [*Berkowitz et al.*, 2002; *Dentz et al.*, 2004]

$$u\tilde{c}(\mathbf{s}, u) - c_0(\mathbf{s}) = -\mathbf{v}^*(u) \cdot \nabla \tilde{c}(\mathbf{s}, u) + \mathbf{D}^*(u) : \nabla\nabla \tilde{c}(\mathbf{s}, u), \quad (24)$$

where we define an advective component

$$\mathbf{v}^*(u) \equiv \int \tilde{\phi}(\mathbf{s}, u)\mathbf{s}\, \mathrm{d}^d s \quad (25)$$

and a dispersive component

$$\mathbf{D}^*(u) \equiv \int \tilde{\phi}(\mathbf{s}, u)\frac{1}{2}\mathbf{ss}\, \mathrm{d}^d s. \quad (26)$$

Note the sum over $\mathbf{s}'$ in (2) is independent of $\mathbf{s}$ in a stationary system; hence we shift the summation variable to obtain (25) and (26).

[46] Observe that (2) does not separate the transport into advective and dispersive components of the motion. Here, too, the separation is only apparent. In fact, the u dependence of (25) and (26) indicates that the nonlocality in time (see discussion following (2)) applies to both the advective and dispersive parts of the motion, where $\mathbf{v}^*(u)$ and $\mathbf{D}^*(u)$ are different spatial moments of the same distribution $\tilde{\phi}(\mathbf{s}, u)$ and are hence connected.

[47] The quality of the approximation (23) depends on the boundedness of the derivatives of $c(\mathbf{s}, t)$ in a sphere of radius s_0 around each point $\mathbf{s}$ and on the boundedness of the spatial moments of $\psi(\mathbf{s}, t)$. The particular formulation in (24) is convenient because we can define terms that are familiar in the context of traditional modeling: the "effective velocity" $\mathbf{v}^*(u)$ and the "dispersion tensor" $\mathbf{D}^*(u)$. Note, however, that both of these quantities are u-dependent and, most significantly, depend fundamentally on $\tilde{\psi}(\mathbf{s}, u)$ (recall (7)). This equation has the form of an ADE generalized to nonlocal time responses as a result of the ensemble average. Solutions of (24) can be obtained for general BCs, and so it is an important starting point for the analysis of transport problems.

[48] A similar expansion to (23), for a single realization, together with a similar Taylor expansion of $w(\mathbf{s}, \mathbf{s}')$, can be substituted into (1) [*Berkowitz et al.*, 2002] to yield

$$\frac{\partial C(\mathbf{s}, t)}{\partial t} = -\mathbf{v}(\mathbf{s}) \cdot \nabla C(\mathbf{s}, t) + \nabla \cdot \nabla[\mathbf{D}(\mathbf{s})C(\mathbf{s}, t)], \quad (27)$$

where $\mathbf{v}(\mathbf{s})$ is the velocity field and $\mathbf{D}(\mathbf{s})$ is the dispersion tensor; both are defined by *Berkowitz et al.* [2002] in terms of spatial moments of $w(\mathbf{s}, \mathbf{s}')$ and hence are dominated by the large values of w. Equation (27) is a generalization of the ADE which we shall use in section 4.

[49] In many instances we can assume that the transition rate probability can be applied in the decoupled form $\psi(\mathbf{s}, t) = p(\mathbf{s})\psi(t)$, where $p(\mathbf{s})$ is the probability distribution of the length of the jumps and $\psi(t)$ is the probability rate for a transition time t between sites. The coupling between the spatial displacement and the transition time involves a velocity that correlates very weakly, if at all, with $|\mathbf{s}|$; for example, the v (magnitude of the velocity) acts as an independent random variable and hence diminishes the coupling between $|\mathbf{s}|$ and t. Especially in an highly heterogeneous system, the rate-limiting steps are due to the statistically rare events such as the encounter with a low-velocity region, and these longer times are not correlated with the spatial displacements. As an example (see section 3.8), in a pore-level picture the variation of spatial displacements (the throat between pores) is rather narrow, while the variation of the time of transit between these pores can be very large because of the extended $1/v$ tail of the velocity histogram. The low v range which limits the transport does not necessarily occur with larger or smaller interpore displacements. The time is not controlled by the length of the displacement but rather the v in the throat. More detailed discussion of the assumption of decoupling is given in section 3.3, where we focus on $\mathcal{L}[\psi(\mathbf{s}, t)] = \tilde{\psi}(u)p(\mathbf{s})$, where $\Sigma_{\mathbf{s}}p(\mathbf{s}) = 1$.

[50] Using the decoupled form $\psi(\mathbf{s}, t) = p(\mathbf{s})\psi(t)$, we can rewrite the generalized velocity and dispersion as [*Berkowitz et al.*, 2002; *Dentz et al.*, 2004; *Cortis et al.*, 2004b]

$$\mathbf{v}^*(u) = \tilde{M}(u)\mathbf{v}_\psi \quad (28a)$$

$$\mathbf{D}^*(u) = \tilde{M}(u)\mathbf{D}_\psi, \quad (28b)$$

where

$$\tilde{M}(u) \equiv \bar{t}u\frac{\tilde{\psi}(u)}{1 - \tilde{\psi}(u)} \quad (29)$$

is a memory function and

$$\mathbf{v}_\psi = \frac{1}{\bar{t}}\int p(\mathbf{s})\mathbf{s}\, \mathrm{d}^d s \quad (30a)$$

$$\mathbf{D}_\psi = \frac{1}{\bar{t}}\int \frac{1}{2}p(\mathbf{s})\mathbf{ss}\, \mathrm{d}^d s \quad (30b)$$

(we have multiplied and divided by a characteristic time $\bar{t}$ [*Berkowitz et al.*, 2002]; see further discussion of $\bar{t}$ in section 3). It will be useful in the following to combine the equations in (30) in the form

$$\mathbf{D}_\psi = \frac{\left|\int p(\mathbf{s})\mathbf{s}\, \mathrm{d}^d s\right|}{\bar{t}} \frac{\int \frac{1}{2} p(\mathbf{s})\mathbf{s}\mathbf{s}\, \mathrm{d}^d s}{\left|\int p(\mathbf{s})\mathbf{s}\, \mathrm{d}^d s\right|} \equiv \alpha_\psi |\mathbf{v}_\psi|, \tag{31}$$

where the tensor

$$\alpha_\psi = \frac{\int \frac{1}{2} p(\mathbf{s})\mathbf{s}\mathbf{s}\, \mathrm{d}^d s}{\left|\int p(\mathbf{s})\mathbf{s}\, \mathrm{d}^d s\right|} \tag{32}$$

has the dimension of length.

[51] As stated in the beginning of this section, the basis for the approximation (23) is the existence of the spatial moments of $\psi(\mathbf{s}, t)$. In the decoupled form this requires $p(\mathbf{s})$ to possess finite moments. The first two of these moments (in (30)) define coefficients of the transport velocity and dispersion. The $p(\mathbf{s})$ can otherwise be a general positive function including one with a long tail (with a large s cutoff). The choice of a Gaussian for $p(\mathbf{s})$ used later in (63) is simply a convenience for the numerical simulation.

[52] The working transport equation for an ensemble-averaged system is (24) with the definitions (28)–(30),

$$u\tilde{c}(\mathbf{s}, u) - c_0(\mathbf{s}) = -\tilde{M}(u)\left[\mathbf{v}_\psi \cdot \nabla \tilde{c}(\mathbf{s}, u) - \mathbf{D}_\psi : \nabla\nabla \tilde{c}(\mathbf{s}, u)\right]. \tag{33}$$

It is important to recognize that the "transport velocity" $\mathbf{v}_\psi$ is distinct from the "average fluid velocity" $\mathbf{v}$, whereas in the classical advection-dispersion picture these velocities are identical. Similarly, the "dispersion" $\mathbf{D}_\psi$ has a different physical interpretation than in the usual ADE definition. The flux $\mathbf{j}$ is defined through $\partial c/\partial t = -\nabla \cdot \mathbf{j}$, so that

$$\tilde{\mathbf{j}}(\mathbf{s}, u) \equiv \tilde{M}(u)\left[\mathbf{v}_\psi \tilde{c}(\mathbf{s}, u) - \mathbf{D}_\psi \cdot \nabla \tilde{c}(\mathbf{s}, u)\right]. \tag{34}$$

[53] A key feature of (33) is that it encompasses various common models, such as multirate and mobile-immobile transport equations [e.g., *Villermaux*, 1974, 1987] for specific examples of $\tilde{M}(u)$ (or $\tilde{\psi}(u)$), together with other simplifications. These issues are discussed in section 7.

[54] Observe also that if desired, one can avoid the use of Laplace transforms by working directly in the time domain, i.e., solving the $\mathcal{L}^{-1}$ of (33):

$$\frac{\partial c(\mathbf{s}, t)}{\partial t} = -\int_0^t M(t - t')\left[\mathbf{v}_\psi \cdot \nabla c(\mathbf{s}, t') - \mathbf{D}_\psi : \nabla\nabla c(\mathbf{s}, t')\right] dt'. \tag{35}$$

[55] The solution of (33), $\tilde{c}(\mathbf{s}, u)$, can now be expressed [see *Dentz et al.*, 2004, Appendix B] in terms of $\tilde{c}_1(\mathbf{s}, u)$, which is the solution of (33) for $\tilde{M}(u) = 1$ and the same BCs as $\tilde{c}(\mathbf{s}, u)$:

$$\tilde{c}(\mathbf{s}, u) = \frac{1}{\tilde{M}(u)}\, \tilde{c}_1\left(\mathbf{s}, \frac{u}{\tilde{M}(u)}\right). \tag{36}$$

As noted in section 2.5, the form for $\psi(t)$ corresponding to the ordered case is an exponential $\psi(t) = W \exp(-Wt)$. Inserting $\tilde{\psi}(u) = W/(W + u)$ into (29), one obtains $\tilde{M}(u) = 1$ with $\bar{t} = W^{-1}$. In this perfectly ordered case, (33) is formally equivalent to the classical ADE. Thus a u dependence of $\tilde{M}(u)$ derives from a range of rates which characterize a disordered system.

[56] We now exhibit the analytic solutions of (33) for different spatial dimensions and simple BCs. (Several other explicit formulas for resident and flux-averaged concentrations involving BCs on the flux (34) at an injection and a control plane are given by *Dentz et al.* [2004]. We assume that $\mathbf{v}_\psi$ is aligned in the one direction of the coordinate system, $\mathbf{v}_{\psi,i} = v_\psi\, \delta_{i1}$, and $D_{\psi,ij}$ is diagonal with $D_{\psi,11} \equiv D_\psi^L$ and $D_{\psi,ii} \equiv D_\psi^T$, $i > 1$, where the superscripts L and T denote longitudinal and transverse, respectively. As a BC in an infinite domain we assume a vanishing $c(\mathbf{s}, t)$ at infinity; as an initial condition, $c(\mathbf{s}, 0) = \delta(\mathbf{s})$. Thus, with the above definitions, the solution to (33) reads for one dimension

$$\tilde{c}(s_1, u) = \frac{\exp\left[-\frac{v_\psi}{2D_\psi}\left(\sqrt{s_1^2 + 4s_1^2 \frac{uD_\psi}{\tilde{M}(u)v_\psi^2}} - s_1\right)\right]}{\tilde{M}(u)\, v_\psi \sqrt{1 + 4\frac{uD_\psi}{\tilde{M}(u)v_\psi^2}}}. \tag{37}$$

In two dimensions we obtain

$$\tilde{c}(\mathbf{s}, u) = \frac{\exp\left(\frac{s_1 v_\psi}{2D_\psi^L}\right)}{2\pi \tilde{M}(u)\sqrt{D_\psi^L D_\psi^T}} \cdot K_0\left(\frac{v_\psi}{2D_\psi^L}\sqrt{s_1^2 + \frac{D_\psi^L}{D_\psi^T}s_2^2}\sqrt{1 + 4\frac{uD_\psi^L}{\tilde{M}(u)v_\psi^2}}\right), \tag{38}$$

where $K_0(z)$ is the modified Bessel function [*Abramowitz and Stegun*, 1970], while for three dimensions the $\mathcal{L}$ of the concentration distribution is given by

$$\tilde{c}(\mathbf{s}, u) = \left(4\pi \tilde{M}(u) D_\psi^T \sqrt{s_1^2 + \frac{D_\psi^L}{D_\psi^T}\left(s_2^2 + s_3^2\right)}\right)^{-1} \cdot \exp\left[-\frac{v_\psi}{2D_\psi^L}\left(\sqrt{s_1^2 + \frac{D_\psi^L}{D_\psi^T}\left(s_2^2 + s_3^2\right)} \cdot \sqrt{1 + 4\frac{uD_\psi^L}{\tilde{M}(u)v_\psi^2}} - s_1\right)\right], \tag{39}$$

where D_ψ^T values in the two directions orthogonal to the principal flow direction are assumed equal. Note that (37)–

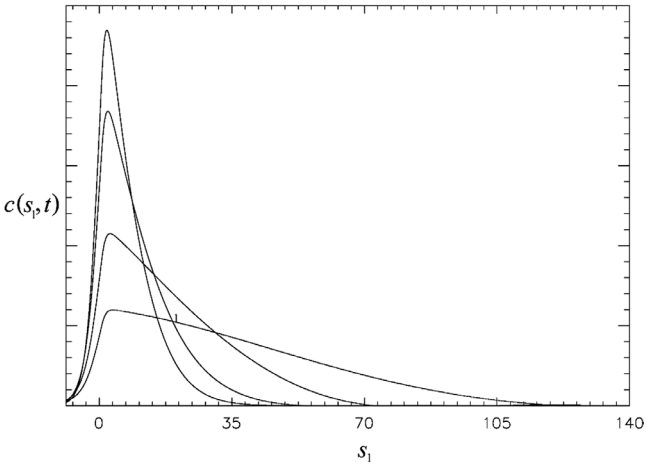

Figure 4. Profiles of the spatial distribution of tracer, $c(s_1, t)$, determined from the CTRW solution for $\beta = 0.5$, at four relative times $\tau = 1$, 2.5, 10, and 37.5. The concentration profile, $c(s_1, t)$, is the vertical average (along the s_2 axis) of $c(\mathbf{s}, t)$ in arbitrary units; particle injection point is at $s_1 = 0$. Reprinted from *Berkowitz and Scher* [1998] (http://dx.doi.org/10.1103/PhysRevE.57.5858). Copyright 1998 by the American Physical Society.

(39), which are valid for arbitrary $\tilde{\psi}(u)$, enable study of the full time behavior of $c(\mathbf{s}, t)$ in one to three dimensions by numerical Laplace inversion ($\mathcal{L}^{-1}$).

3. MODELS OF $\psi(\mathbf{s}, t)$ AND APPLICATIONS TO LABORATORY AND FIELD EXPERIMENTS

[57] We have discussed in section 2 how $\psi(\mathbf{s}, t)$ is the heart of the CTRW. We noted in section 2.1 that particle transitions are influenced by a wide range of mechanisms, and in section 2.5 we related these transitions to $\psi(\mathbf{s}, t)$. Thus $\psi(\mathbf{s}, t)$ can, in principle, be defined to account for advective, diffusive, and/or dispersive displacement over a wide range of timescales. Specifically, $\psi(\mathbf{s}, t)$ can account for transport in both fully and partially saturated media in domains containing flowing and stagnant (mobile and immobile, fracture, and host rock) zones. Moreover, one can define a $\psi(\mathbf{s}, t)$ that includes effects of sorption [e.g., *Margolin et al.*, 2003]. We now consider in this section specification of $\psi(\mathbf{s}, t)$ and application of CTRW to a variety of laboratory and field measurements.

3.1. Definition of $\psi(\mathbf{s}, t)$ for a Fracture System

[58] A relatively simple physical example that allows specification of $\psi(\mathbf{s}, t)$ is transport in a "random" fracture network (RFN) [*Berkowitz and Scher*, 1997, 1998; *Scher et al.*, 2002a]. Such discrete fracture conceptualizations have proven useful in a variety of situations [e.g., *National Research Council*, 1996]. We picture a RFN as an extended set of intersecting channels, tubes, or discs filled with a fluid in steady flow. We then characterize the RFN by a distribution of fracture fragments (the lengths between intersections) $f(s)$ and by a velocity histogram (the distribution of velocities in each tube fragment) $\Phi(\xi, \theta)$, where $\xi = 1/v$ (with v the magnitude of the velocity) and θ is the angle orientation of the fragment with respect to the pressure gradient ($|\mathbf{s}| = s$, $s_x = s\cos\theta$). The transition time is simply the time of transit in each fragment. Thus

$$\psi(\mathbf{s}, t) = C_n \Phi(\xi, \theta) f(s), \tag{40}$$

where C_n is a normalization constant and $t = s\xi$.

[59] A simplified version of the function we used to fit simulation data [*Berkowitz and Scher*, 1997, 1998; *Scher et al.*, 2002a] is

$$\Phi(\xi, \theta) \propto \xi^{-1-\beta} \exp\left[-\xi_o / \left(\xi \cos^2 \frac{\theta}{2}\right)\right], \tag{41}$$

where β characterizes the width of the distribution at low velocities ($\beta \leq 0.9$ for the simulations and $\beta \simeq \frac{1}{2}$ for the data given by *Scher et al.* [2002a]) and $f(s) \propto s^{1/2} e^{-s/s_o}$. The parameters in the distributions can be used to form dimensionless variables characteristic of this problem: the time $\rightarrow \tau \equiv t/\bar{t}$, the Laplace variable $u \rightarrow \mu \equiv \bar{t}u$, where $\bar{t} \equiv \frac{1}{2} s_o \xi_o$, and the displacement $\rightarrow \rho \equiv s/s_o$. Typically, $\bar{t}$ represents the median time for one transition. In the applications we will be interested in particle transport after many transitions, i.e., $\tau \gg 1$, hence the range $\mu \ll 1$.

[60] The general solution for the plume (8) with periodic BCs requires both the $\mathcal{F}$ and $\mathcal{L}$ of $\psi(\mathbf{s}, t)$. The $\mathcal{L}[\psi(\mathbf{s}, t)]$ can be determined analytically:

$$\tilde{\psi}(\mathbf{s}, u) = C_n' \varrho^{1/2} e^{-\varrho} \varrho \left[\cos^\beta \frac{\theta}{2}\right] \left(\sqrt{2\varrho\mu}\right)^\beta K_\beta \left(2\sqrt{2\varrho\mu} / \cos\frac{\theta}{2}\right), \tag{42}$$

where $K_\beta(x)$ is the modified Bessel function of order β. The $\mathcal{F}(\tilde{\psi}(\mathbf{s}, u))$ must be evaluated numerically, and finally, the plume $c(\mathbf{s}, t)$ is determined by the inverse $\mathcal{F}$ and $\mathcal{L}$ of (8). We show in Figure 4 a plot of the vertical average of the plume $c(x, t)$ for four large values of τ and for $\beta = \frac{1}{2}$. The plume is highly non-Gaussian, a characteristic of anomalous transport that will be detailed in section 3.2. This plume shape was seen clearly in simulations of particle migration in a RFN, as shown in Figure 5. The β of the simulations shown in Figure 5 corresponds to a value of $\beta \approx 0.8$. For values of $\frac{1}{2} < \beta < 1$ the peak moves (as opposed to the curves shown in Figure 4), and the mean position moves faster than the mean position of the plume in Figure 4. Hence the comparison between Figures 4 and 5 is qualitative.

3.2. Asymptotic Behaviors

[61] Recall that as discussed in section 2.5, the term "asymptotic" is used to refer to the long time, or small μ, behavior in terms of the particle transitions. The source of the non-Gaussian behaviors discussed previously can be seen from (42) in the limit $\mu \rightarrow 0$, i.e., $\tau \rightarrow \infty$:

$$\Lambda(0, u) \simeq 1 - \sqrt{2\pi}\, \frac{16}{5}\, \mu^{1/2} - 7\mu(\ln\mu) + \mathcal{O}(\mu). \tag{43}$$

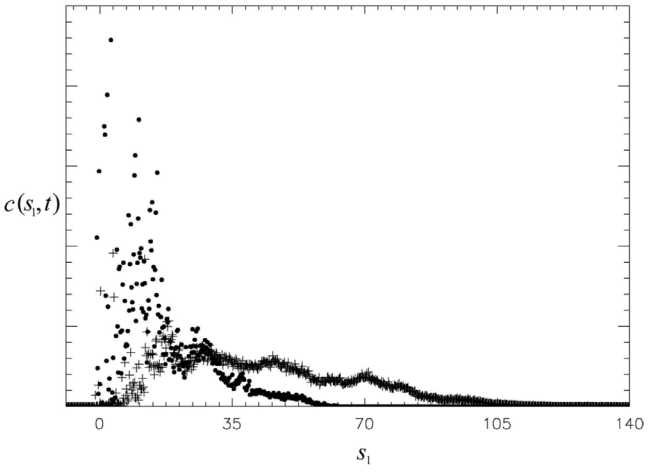

Figure 5. Profile of the spatial distribution of tracer at two relative times ($\tau = 1$, dots, and $\tau = 2.5$, crosses), determined from numerical simulations of particle transport in a two-dimensional, random fracture network. The concentration profile, $c(s_1, t)$, is the vertical average (along the s_2 axis) of $c(\mathbf{s}, t)$ in arbitrary units; particle injection point is at $s_1 = 0$. Reprinted from *Berkowitz and Scher* [1998] (http://dx.doi.org/10.1103/PhysRevE.57.5858). Copyright 1998 by the American Physical Society.

The first temporal moment (mean time) of $\Sigma_{\mathbf{s}}\psi(\mathbf{s}, t)$ ($\equiv \psi(t)$, see (5)) $\hat{t}$ is equal to the derivative $d\Lambda(0, u)/du|_{u=0}$. For (43), $\hat{t}$ diverges, which is also the case for

$$\tilde{\psi}(u) \overset{\mu\to 0}{\longrightarrow} 1 - c_\beta \mu^\beta, \qquad 0 < \beta < 1, \tag{44}$$

where c_β is a constant. The exact low μ expansion of (42) is carried out in Appendix D.

[62] The behavior in (44) originates in the form of the low-velocity tail $\xi^{-1-\beta}$ in (40) and (41), which yields

$$\psi(t) \propto \tau^{-1-\beta}, \qquad \tau \longrightarrow \infty. \tag{45}$$

Hence, if there is a sufficient range of a slow (i.e., algebraic) t dependence such as in (45), one has either a divergent $\hat{t}$ or one that is larger than the duration of the observation. This leads to a transport that is effectively independent of a characteristic time. In Appendix E we present a general derivation of (44) from (45).

[63] For the large t dependence in (45) we have for the mean displacement $\bar{\ell}(t)$ and standard deviation $\bar{\sigma}(t)$ of the plume $c(\mathbf{s}, t)$ that [*Shlesinger*, 1974]

$$\bar{\ell}(t) \propto t^\beta, \quad \bar{\sigma}(t) \propto t^\beta, \quad t \longrightarrow \infty, \quad 0 < \beta < 1. \tag{46}$$

This behavior is in sharp contrast to Gaussian models, where $\bar{\ell}(t) \propto t$ and $\bar{\sigma}(t) \propto t^{1/2}$ as an outcome of the central limit theorem; in this case the position of the peak of the distribution coincides with $\bar{\ell}(t)$. It is worth noting that the range $1 < \beta < 2$ is also a non-Fickian regime for although $\bar{\ell}(t) \propto t$, the standard deviation is $\bar{\sigma}(t) \propto t^{(3-\beta)/2}$, which is not the Gaussian case.

[64] We note here, for completeness, that to quantify diffusion-only systems, we can set $\mathbf{v}_\psi = 0$ in (33) and use a $\psi(t)$, defined for $0 < \beta < 1$, based on pure diffusion. In this case, $\bar{\ell}(t) = 0$ and $\bar{\sigma}(t) \propto t^{\beta/2}$ [*Shlesinger*, 1974], as opposed to t^β in the advective transport case given by (46). Recent detailed developments considering anomalous diffusion are described elsewhere [*Metzler and Klafter*, 2004; *Hornung et al.*, 2005].

[65] It is important to recognize that because of non-locality the use of time-dependent coefficients in a conventional pde, e.g., as was done with the ADE [see *Dagan and Neuman*, 1997], is conceptually incorrect; rather, a nonlocal pde of the form in (35) is required. If the t dependence of $\psi(t)$ follows the form in (45) over the duration of the observations then the transport will be anomalous.

[66] Use of the asymptotic form (44) simplifies the calculation of $f_B(t)$ in (12),

$$f_B(\tau) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\mu\tau} \exp\left(-c_\beta L \mu^\beta / \bar{l}\right) d\mu, \tag{47}$$

for $\sigma > 0$ and $0 < \beta < 1$ [*Montroll and Scher*, 1973; *Scher and Montroll*, 1975]. The $f_B(t)$ in (47), valid in the asymptotic limit, is equal to one of the Fox H functions [*Metzler and Klafter*, 2000].

[67] It is expedient to work directly with the $\mathcal{L}^{-1}$ in (47), which can be evaluated by using highly accurate, approximate methods. The results for (47) and the integral of $f_B(t)$ (the flux at $s_1 = L$ due to a constant initial flux at $s_1 = 0$) are shown in Figure 6. Instead of a Fickian solution, there is an entire family of distributions, one for each value of β ($0 < \beta < 1$). For $\beta \leq 0.5$ the dispersion is very large. As $\beta \to 1$, the long time tail diminishes sharply. For $1 < \beta < 2$ one adds to the μ^β term in (47) a term proportional to μ, and for $\beta > 2$ one adds a term proportional to μ^2. As $\beta \to 2$, the dispersion becomes Fickian.

[68] The parameter β is, in effect, a measure of the dispersion in conjunction with $\mathbf{D}_\psi$. It determines the characteristic shape of the BTC and not just the second moment of the distribution. We have used these results to fit

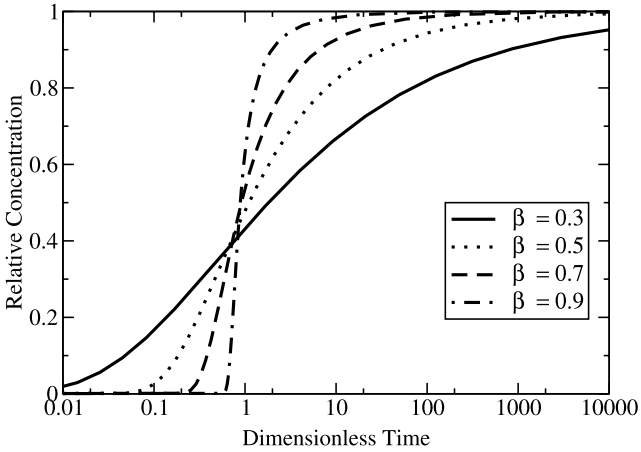

Figure 6. Semilog plot showing a range of cumulative $f_B(t)$ curves, relative concentration versus dimensionless time, for several values of β. Reprinted from *Kosakowski et al.* [2001]. Copyright 2001, with permission from Elsevier.

laboratory and field data of both porous media and fracture networks, as discussed in sections 3.4 and 3.6.

3.3. Coupled and Decoupled $\psi(\mathbf{s}, t)$

[69] In section 2.6 we introduced the "decoupled" form of $\psi(\mathbf{s}, t)$, i.e., $\tilde{\psi}(\mathbf{s}, u) \approx p(\mathbf{s})\tilde{\psi}(u)$, where $p(\mathbf{s})$ denotes the transition length distribution. In the RFN problem we developed a "coupled" form in (42). We now examine the nature of the approximation of a decoupled form in comparison to the coupled one in the asymptotic limit $\mu \to 0$.

[70] Expanding the Bessel function as carried out in Appendix D, the small μ ($\equiv \bar{t}u$) behavior (for $\beta < 1$) of (42) takes the form

$$\tilde{\psi}(\mathbf{s}, u) \approx p_1(\mathbf{s}) - p_2(\mathbf{s})\mu^{\beta}$$
$$\tilde{\psi}(u) \approx 1 - \bar{c}_{\beta}\mu^{\beta}, \tag{48}$$

where $p_1(\mathbf{s}) \equiv \tilde{\psi}(\mathbf{s}, 0)$ and the constant $\bar{c}_{\beta} = \Sigma_{\mathbf{s}} p_2(\mathbf{s})$. The exact forms of $p_1(\mathbf{s})$ and $p_2(\mathbf{s})$ for (42) appear in (D6). We substitute the forms in (48) into (25), where $\tilde{\phi}(\mathbf{s}, u)$ is given in (7), to obtain

$$\mathbf{v}^*(u) \approx \left(\bar{c}_{\beta}\right)^{-1} u\left[\bar{\mathbf{s}}\mu^{-\beta} - \bar{\mathbf{s}}'\right] \approx \left(\bar{c}_{\beta}\right)^{-1}\mu^{1-\beta}\mathbf{v}_{\psi}, \tag{49}$$

with $\bar{\mathbf{s}} = \Sigma_{\mathbf{s}}\mathbf{s}p_1(\mathbf{s})$, $\bar{\mathbf{s}}' = \Sigma_{\mathbf{s}}\mathbf{s}p_2(\mathbf{s})$, and $\mathbf{v}_{\psi} \equiv \bar{\mathbf{s}}/\bar{t}$. Notice that to leading order in the small μ limit we drop the higher-order term in (49). The singular behavior of (49) is controlled by the denominator in (7). In the asymptotic limit these statements are true for a general $\tilde{\psi}(\mathbf{s}, u)$.

[71] Consider now the decoupled joint pdf, $\tilde{\psi}(\mathbf{s}, u) = p(\mathbf{s})\tilde{\psi}(u) \approx p(\mathbf{s})[1 - c_{\beta}(\bar{t}'u)^{\beta}]$, with c_{β} a constant. We find that we can reproduce the result in (49) with this decoupled form if we set the first moment of $p(\mathbf{s})$ equal to that of $p_1(\mathbf{s})$, recalling also that $\mu \equiv \bar{t}u$. In other words, in the small μ limit we see that the decoupling of $\psi(\mathbf{s}, t)$ simply modifies $\bar{\mathbf{s}}$ and $\bar{t}$, with $\bar{t}' \to \bar{t}(\bar{c}_{\beta}/c_{\beta})^{1/\beta}$.

[72] Hence, because the moment $\bar{\mathbf{s}}$ is chosen to correspond to the mean size of the heterogeneous distribution (for the ensemble-averaged system) and the $\bar{t}$ is a characteristic median time for a transition, we can justify, in most situations, the use of the decoupled form $\psi(\mathbf{s}, t) = p(\mathbf{s})\psi(t)$ if the low μ behavior of the coupled $\psi(\mathbf{s}, t)$ has the form (48). Recall also the discussion on the lack of correlation between $\mathbf{s}$ and v in the paragraph preceding the one containing (28).

3.4. Laboratory and Field Experiments

[73] The solutions given at the end of section 3.2, and demonstrated in Figure 6, are a key result, convenient for describing a wide range of transport experiments. These CTRW solutions are based on an asymptotic (algebraic) form of $\psi(t)$ given by (44). In sections 3.4.1 and 3.4.2 we exhibit the application of these solutions to measured BTCs from both field- and laboratory-scale experiments.

[74] It is notable that similar CTRW solutions have been implemented in a variety of studies. In addition to those described throughout sections 3.4–3.8 and section 4, CTRW has successfully quantified tracer transport in partially saturated porous medium columns [*Bromly and Hinz*, 2004] and in single fissures [*Jiménez-Hornero et al.*, 2005], transport of sorbing tracers in porous medium columns [*Hatano and Hatano*, 1998], and transport of colloids in a shear zone [*Kosakowski*, 2004], as well as numerically simulated transport of tracers in faults [*O'Brien et al.*, 2003a, 2003b] and in heterogeneous media [*Di Donato et al.*, 2003; *Bijeljic and Blunt*, 2006].

3.4.1. Field Experiments

[75] We consider a field application that follows naturally from the discussion in section 3 on transport in a RFN. *Sidle et al.* [1998] report on a field-scale tracer test in a fractured till of glacial origin. The matrix permeability of the till is low and therefore water flow is attributed mainly to the fractures. The system is conceptualized as consisting of layers that become less densely fractured and perturbed with depth. However, variation in the hydraulic properties of these layers is unknown, as is the degree of fracture connectivity and the hydraulic properties of individual fractures.

[76] In the field experiment a block of till was excavated and horizontal multiple-port samplers were installed at depths of 2.5 m (relatively heavily fractured) and 4.0 m (less fractured). An infiltration basin installed near the ground surface was used to establish steady state inflow conditions. A chloride tracer was then infiltrated for 7 days, during which water samples were collected and BTCs were recorded. A detailed description of the site and the tracer test is given by *Sidle et al.* [1998].

[77] Typical BTCs for the sampling locations at the 2.5 and 4.0 m depths (screens, B1 and F4, respectively) are shown in Figure 7. These BTCs were analyzed by *Sidle et al.* [1998] using best fits by the one-dimensional ADE, referred to as equivalent porous medium (EPM) model, and by a discrete fracture model (DFM) based on a parallel fracture conceptualization. The DFM, which assumes advective transport in the fractures, with diffusion into and within the host matrix, is in some sense more flexible than the EPM, because it allows also for prescription of combinations of fracture aperture and fracture spacing. However, as seen from Figure 7, neither of these approaches captures the full evolution of the measured BTCs. Moreover, in both cases, estimated model parameter values (e.g., water velocity, dispersivity, fracture aperture, and fracture spacing) for the various BTCs varied over more than 1 order of magnitude, with differences of up to 2 orders of magnitude in estimated water velocities between the EPM and DFM models.

[78] The heterogeneous nature of the medium and the non-Fickian nature of the transport reflected in these BTCs suggest consideration of the CTRW conceptual picture and framework. As discussed in detail by *Kosakowski et al.* [2001], such a treatment is indeed effective. Figure 7 displays fits to the BTCs using the CTRW solution (47), which is based on the asymptotic form of $\psi(t)$ given by (45). The estimated values of β that fit the BTCs are significantly smaller than unity, indicating a clear non-Fickian transport behavior. As discussed in section 3.2, smaller values of β are indicative of more dispersive, heterogeneous systems.

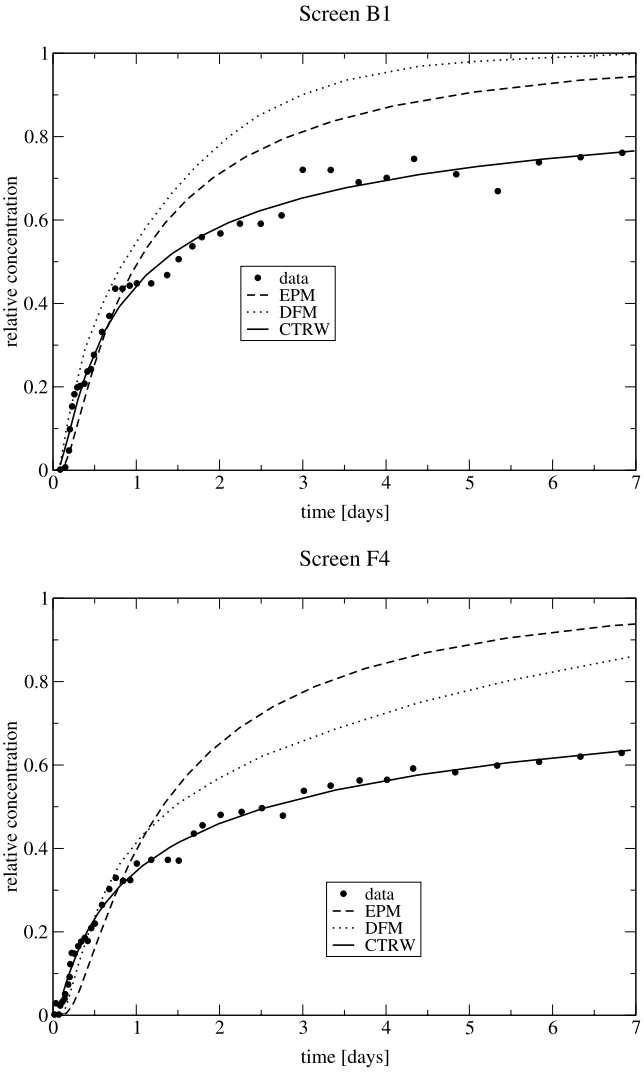

Figure 7. Best fit theoretical BTCs based on solutions for the equivalent porous medium (EPM = ADE), the discrete fracture model (DFM), and the CTRW (with $\beta = 0.49$ for screen B1 and $\beta = 0.39$ for screen F4). Dots indicate measured values. Screens B1 and F4 were located at depths of 2.5 and 4 m, respectively. Reprinted from *Kosakowski et al.* [2001]. Copyright 2001, with permission from Elsevier.

Accordingly, the smaller β value for screen F4, relative to that for the screen B1 BTC, is consistent with the initial geological mapping of *Sidle et al.* [1998], which indicated sparser, more heterogeneous fracturing at the greater depth.

[79] Another application of CTRW solutions has been to analysis of measurements from a large-scale field study in a heterogeneous alluvial aquifer at the Columbus Air Force Base (Mississippi), otherwise known as the "Macrodispersion Experiment" or "MADE" [*Adams and Gelhar*, 1992]. Bromide was injected as a pulse and traced over a 20 month period by sampling from an extensive three-dimensional well network. The tracer plume that evolved was remarkably asymmetric and cannot be described by classical Gaussian models.

[80] Figure 8 displays a direct comparison of the advance of the measured tracer plume and the CTRW solution, using $\beta = 1/2$, at two times. The CTRW solution was fit approximately to the data at $t = 49$ days. Because of the noise in the measurements (the sharp peaks of the tracer data are largely an artifact of the coarse sampling step sizes) more precise fitting is not justified. The comparison to the data at $t = 370$ days represents a prediction of the CTRW solution, using the same parameters as fit previously and just increasing the time. Clearly, the CTRW formulation quantitatively captures the plume behavior. Full details of the analysis of these data are given by *Berkowitz and Scher* [1998].

[81] The connection between the non-Fickian transport measured at the Columbus field site and the CTRW framework can be strengthened further by considering the time dependence of the spatial mean and standard deviation of the field plumes. Analyzing the behavior of the measured mean displacement $\bar{\ell}(t)$ and standard deviation $\bar{\sigma}(t)$, along the longitudinal principal axis of the tracer plume, it was found [*Berkowitz and Scher*, 1998] that they both scale as $\sim t^{0.6\pm0.1}$ in accordance with (46).

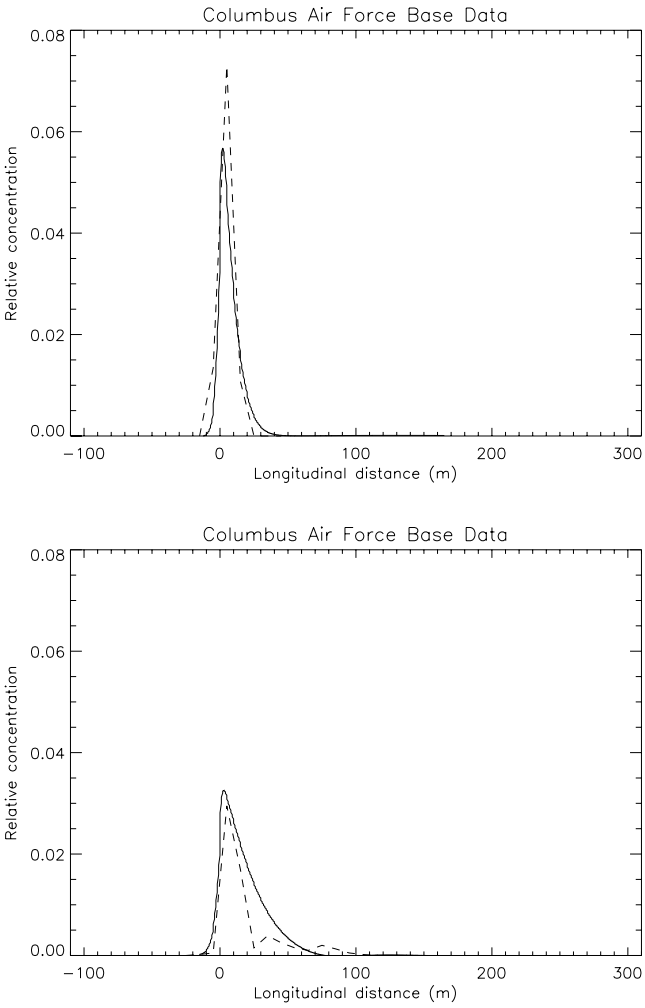

Figure 8. Comparison of the advance of the measured tracer plume (dashed line, Columbus Air Force Base field data from *Adams and Gelhar* [1992]) and the CTRW solution with $\beta = 0.5$ (solid line): (top) $t = 49$ and (bottom) $t = 370$ days. Reprinted with permission from *Berkowitz and Scher* [1998] (http://dx.doi.org/10.1103/PhysRevE.57.5858). Copyright 1998 by the American Physical Society.

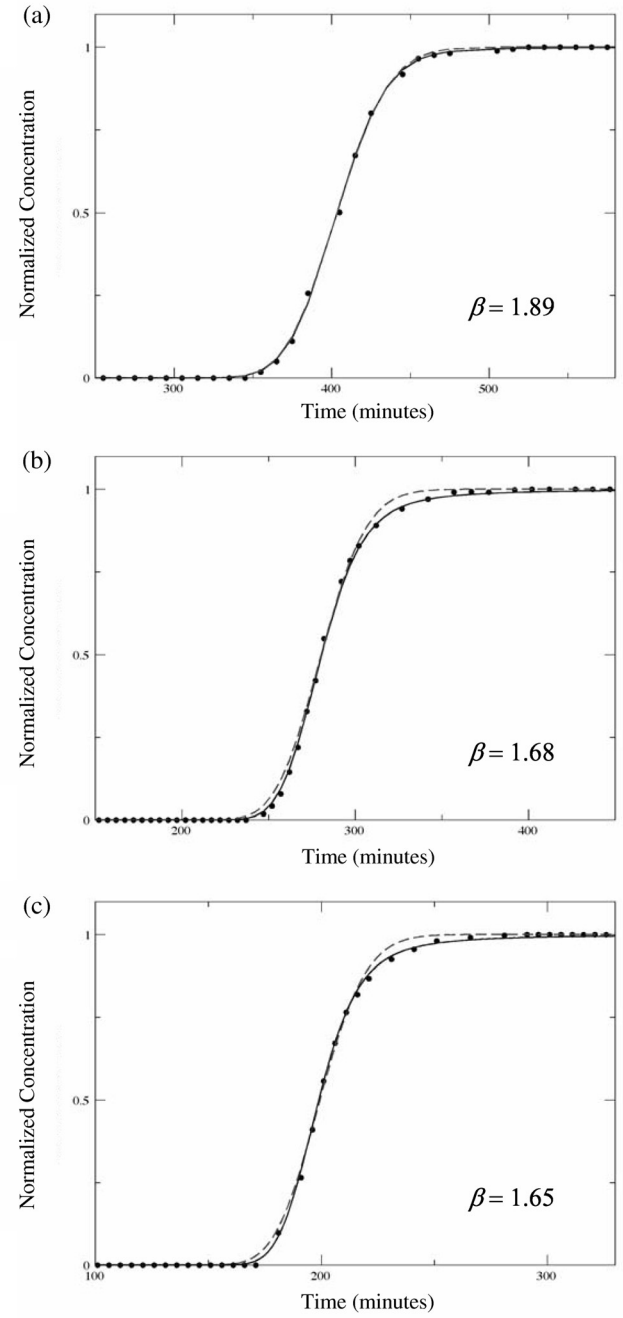

Figure 9. Measured BTCs with fitted ADE (dashed lines) and cumulative $f_B(t)$ (solid lines) solutions for the homogeneous sand pack shown in Figure 1. Flow rates for each experiment were (a) 36, (b) 53, and (c) 74 mL/min. Values of β are indicated for each fit. Corresponding values of dispersion D for the ADE fits are 0.037, 0.072, and 0.120 cm^2/min (Figures 9a–9c, respectively). Reprinted from *Levy and Berkowitz* [2003]. Copyright 2003, with permission from Elsevier.

[82] The CTRW plume shape precludes the need to assume that the bromide was injected initially into a low-permeability region. Furthermore, the long forward tail, which eludes detection, can be the reason for the reported loss of tracer mass. In spite of the uncertainty surrounding this data set the CTRW analysis effectively accounts for the time-dependent plume pattern.

3.4.2. Laboratory Experiments

[83] Over the last 5 decades many dispersion experiments have been performed in small-scale columns and flow cells with dimensions of up to several tens of centimeters in length. As mentioned in section 1, measurements of transport in a variety of porous medium packing arrangements have formed the basis for development of modeling transport behavior, but an accounting of the heterogeneities that occur even at small scales has not been sufficient. Indeed, the non-Fickian transport behavior shown in Figure 1 clearly demonstrates fundamental properties that must be treated in a transport theory.

[84] *Silliman and Simpson* [1987] provide a specific set of measurements that demonstrate non-Fickian transport due to small-scale heterogeneities in porous sand packs. The reported BTCs were analyzed subsequently by *Berkowitz et al.* [2000], using a CTRW solution identical to that used in section 3.4.1. Three BTCs measured at increasing distances from the inlet boundary were examined. In each case the transport behavior was captured, with $\beta \approx 0.87 \pm 0.01$. More detailed BTC measurements are reported and analyzed by *Levy and Berkowitz* [2003].

[85] In particular, *Levy and Berkowitz* [2003] measured BTCs in the "homogeneous," uniform sand pack system shown in Figure 1, finding subtle yet measurably significant differences from Fickian behavior and transport described by the ADE. Another key feature of the experiments is the measurement of BTCs for different flow rates (water velocities). Figure 9 shows fitted values of β for different flow rates, contrasting fits of CTRW and ADE, again using the CTRW solution (47). Two specific features are apparent in Figure 9. First, $\beta < 2$, indicating clear non-Fickian transport behavior even in a "homogeneous" domain. Second, it is clear that β decreases as the flow rate (water velocity) increases. A change in the timescale in an heterogeneous domain has a strong effect on the non-Fickian nature of the transport. We address this issue further in the last paragraph of section 3.5 and in section 6.1.2.

[86] Again, it is important to emphasize that the $\psi(t)$ is the intrinsic characterization of the transport. In Figure 3 one can observe that the interplay between the duration of the experiment and the extent of heterogeneity determines the effective position of the time dependence of $\psi(t)$, e.g., the value of β. At sufficiently long times the $\psi(t)$ exhibits a departure from a power law tail, and one can expect an evolution to normal transport in this time interval (see section 3.7). Hence the β is a highly useful parameter determining the transport over an appropriate time duration.

[87] Similar experiments were carried out by *Levy and Berkowitz* [2003] for a "uniformly heterogeneous" sand packing configuration. Contaminant plume behavior visualized by the dye tracer in Figure 10 shows retention of tracer near the inlet of the flow cell over long times, which will lead to late time tailing in BTCs. Fits with CTRW solutions and behavior of the β parameter with increasing flow rates are similar to those shown in Figure 9. To further illustrate

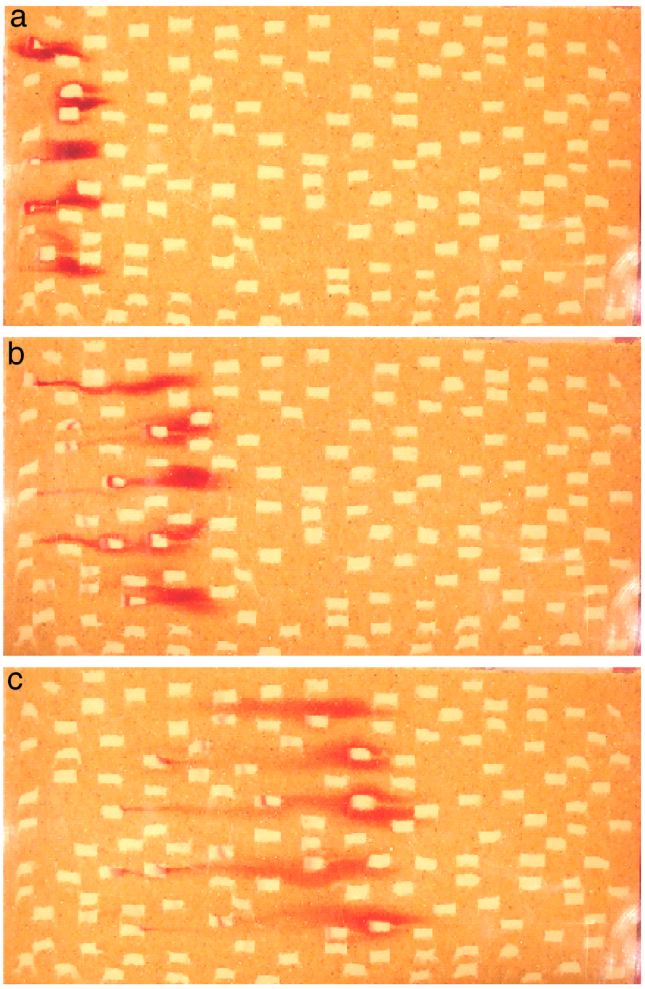

Figure 10. Photographs of a randomly heterogeneous, saturated sand pack with five dye tracer point injections being transported, under constant flow of 65 mL/min, from left to right at (a) $t = 16$ min, (b) $t = 49$ min and (c) $t = 115$ min after injection. Reprinted from *Levy and Berkowitz* [2003]. Copyright 2003, with permission from Elsevier.

the character of the non-Fickian transport behavior that arises, Figure 11 shows spatial concentration profiles that arise for $\beta > 1$, generated using the asymptotic form of $\psi(t)$ given by (45) (but with a modified form of $\tilde{\psi}(u) \simeq 1 - \hat{t}\mu + c\mu^{\beta}$, $1 < \beta < 2$). Recall that Figure 4 shows spatial concentration profiles for similar solutions with $\beta < 1$. In section 3.7 we shall show two-dimensional profiles valid for a full time range.

[88] We now return to consider the measured BTCs of *Scheidegger* [1959] presented in section 1 and subsequent experiments, which could not be captured by the ADE. These classical experiments were reexamined by *Cortis and Berkowitz* [2004] in the context of CTRW. *Scheidegger* [1959] measured BTCs in Berea sandstone cores 30 inches (76.2 cm) long and 2 inches (5.08 cm) in diameter. The cores were fully saturated with tracer and subsequently flushed with clean liquid. The experimental setup ensured a high degree of control of BCs and measurements. Figure 12 presents the BTC data given by *Scheidegger* [1959]. Also shown are best fit curves obtained from solutions of the CTRW and ADE formulations. *Scheidegger* [1959] observed that there were systematic deviations between the computed ADE values and the observed ones: The calculated values were below the measured ones for small and large times, whereas they were above the measured ones for intermediate times. The CTRW solution (using a variant of the $\tilde{\psi}(u)$ in (44)) captures the full evolution of the transport behavior. Note, in particular, that both the (non-Fickian) early and late time portions of the curve are captured well. In contrast, parameters can be chosen to

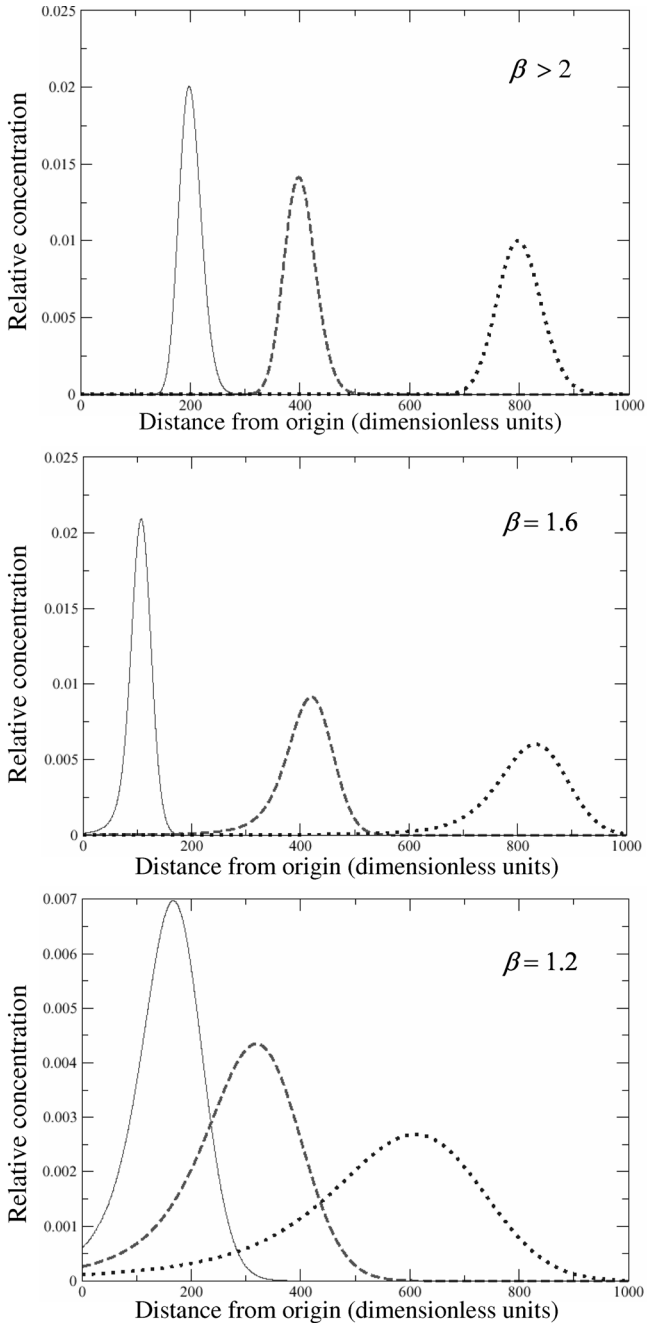

Figure 11. Spatial profiles of tracer plumes at $\beta > 2$, $\beta = 1.6$, and $\beta = 1.2$. The solid, dashed, and dotted curves in each plot represent early, intermediate, and late dimensionless times, respectively. Reprinted from *Levy and Berkowitz* [2003]. Copyright 2003, with permission from Elsevier.

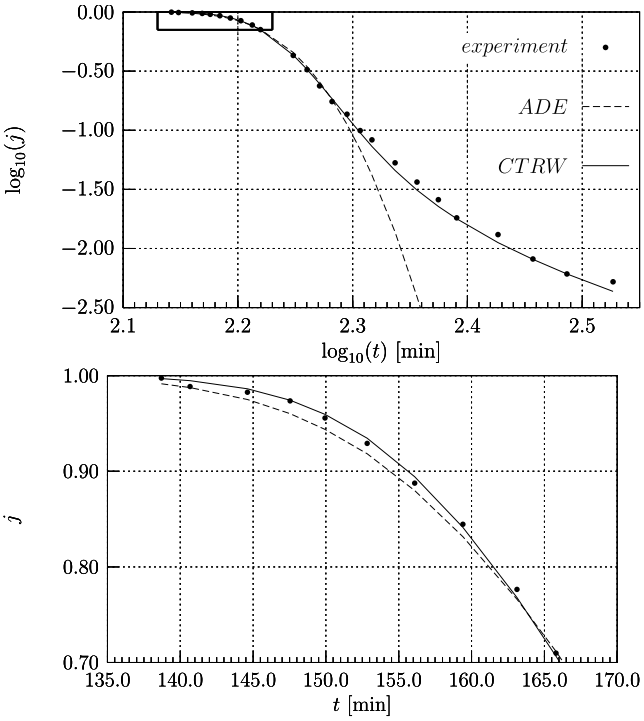

Figure 12. Measured BTC with CTRW/ADE fits for a Berea sandstone core, using a truncated power law function (61) with $\beta = 1.59$. Here the quantity j represents the normalized, flux-averaged concentration. (top) Complete BTC. (bottom) Region identified by the bold-framed rectangle in Figure 12 top plot. Note the difference in scale units between the plots. Column length equals 0.762 m. Porosity $n = 0.204$. Flux $q = 1.73$ cm^3/min [after *Scheidegger*, 1959, Figure 2]. Dashed line is the best ADE model fit. Solid line is the best CTRW fit. After *Cortis and Berkowitz* [2004]. Reproduced with permission.

enable the ADE to fit, at best, only the transport of the center of mass of the tracer plume.

[89] *Cortis and Berkowitz* [2004] used a similar analysis to reexamine three "typical" BTCs from a series of miscible displacement experiments in both saturated and unsaturated soils, presented by *Nielsen and Biggar* [1962]. The samples were packed in columns 30 cm long. Two BTCs correspond to experiments using columns filled with Aiken clay loam 0.23–0.50 mm aggregates, in saturated conditions, with imposed microscopic velocities of 3.40 cm/h and 0.058 cm/h, respectively. The third BTC corresponds to a partially saturated (volumetric water content $\theta = 0.27$) Oakley sand with an imposed microscopic velocity equal to 1.03 cm/h. Similar to *Scheidegger* [1959], *Nielsen and Biggar* [1962] also reported systematic deviations of the calculated values using the ADE from the experimental data, all of which display non-Fickian transport behavior. As before, the CTRW solutions capture the BTC behavior more completely than those provided by the ADE, as seen in Figure 13.

3.5. Local Fluctuations in Porous Media

[90] We have shown in section 3.4.2 how tracer transport in even "homogeneous" porous media can be non-Fickian. We further examine the subtle effects of pore-scale disorder in the case of a series of uniformly packed arrangements of small spheres of uniform diameter. Clearly, it is not possible to realize a periodic arrangement of spheres from a practical point of view, as some disorder will appear evident at a length scale, say, of a few sphere diameters. The key issue is to recognize that short-range "geometrical" heterogeneity in a porous medium leads also to short-range spatial variability in the microscopic velocity field $\mathbf{v}(\mathbf{y})$ of the liquid that carries the tracer. The complex nature of the paths traveled by the tracer in this type of porous medium is clearly governed by the fluctuations of the Stokes velocity field $\mathbf{v}(\mathbf{y})$ around its volume average value $\langle \mathbf{v}(\mathbf{y}) \rangle$ because of these local heterogeneities in the pore volume.

[91] We consider local averages of the key quantities and retain a statistical characterization of these fluctuations. We "coarse grain" the porous medium by averaging over a V_{cell}; that is, we shift the length scale from the interpore distance to the V_{cell} radius r. The spatial autocorrelation of the local Stokes velocity field can be expressed as

$$\langle v(y)\, v(y+s) \rangle = \langle v(y)^2 \rangle \exp\left(-\frac{s}{\delta}\right), \tag{50}$$

where δ is the correlation length of the fluctuations and the average is over V_{cell} (with $\delta \leq r$). We simplify the problem

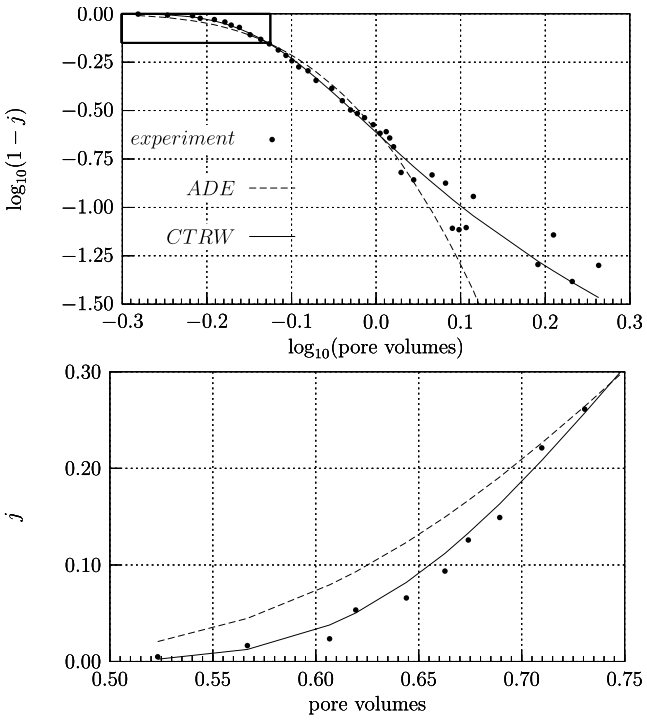

Figure 13. Measured BTC with CTRW/ADE fits for Aiken clay loam 0.25–0.50 mm aggregates, using a truncated power law function (61) with $\beta = 1.29$. Here the quantity j represents the normalized, flux-averaged concentration. (top) Complete BTC. (bottom) Region identified by the bold-framed rectangle in the Figure 13 top plot. Note the difference in scale units between the plots. Column length equals 30 cm. Velocity $v = 0.058$ cm/h [after *Nielsen and Biggar*, 1962, Figure 9], dots). Dashed line is the best ADE model fit. Solid line is the best CTRW fit. After *Cortis and Berkowitz* [2004]. Reproduced with permission.

by considering the autocorrelation in the bias direction only and hence use scalar notation.

[92] We assume that the local transition rate for the tracer is due to advection in this Stokes field, as well as diffusion:

$$w(y+s) = \frac{a}{8}\frac{v(y+s)}{\bar{l}} + \frac{b}{6\bar{t}}\exp\left(-\frac{s}{\lambda}\right), \tag{51}$$

with λ being a characteristic diffusion length, $\bar{l} = \Sigma_s p(s)\, s$ being a characteristic pore length, and $\bar{t}$ being a characteristic "relaxation time for diffusion." The constants a and b are fitted to the experimental data. The transition rate in (51) also fluctuates and therefore must be characterized statistically.

[93] *Cortis et al.* [2004a] derive an effective transition rate, which involves (50) and (51), from a coarse-grained master equation (ME) for the flux-averaged concentration. The effective rate is of the form (20) with a velocity-dependent prefactor

$$w_m \propto \eta \equiv a\frac{\langle v(y)\rangle}{v_\psi} + b, \tag{52}$$

where the transport velocity $v_\psi \equiv \bar{l}/\bar{t}$, $r_o \to \delta$, and $\lambda \sim \delta$. The transitions of the form (20) are taken between randomly distributed volumes, V_{cell}, for this case of mild disorder (valid for $\eta \sim 1$). Thus the $\psi(\tau)$ is the same as (21) with $\tau \equiv (t/\bar{t})\eta/6$ and $\zeta \to \eta$; as a result, we denote (21) as the η function. The result depends on a natural ansatz that the radius of V_{cell} is velocity-dependent, as the range of the transitions depend on v, which we choose as $r \equiv 6\delta/\eta$ [*Cortis et al.*, 2004a].

[94] The BTC measurements were carried out on homogeneous sand columns [*Cortis et al.*, 2004a], and surprisingly, we observed distinct effects due to these small-scale grain-size-level heterogeneities. A series of tracer breakthrough experiments was performed in a one-dimensional flow field on uniformly packed columns of two different lengths. Each column contained either glass beads or well-rounded quartz sand in one of three (fine, medium, or coarse) average grain sizes.

[95] Figure 14 shows a typical BTC for a short column (L = 20 cm) filled with fine sand. It can be observed that while the mean arrival time is matched approximatively by the fitted ADE model, the predicted early and late arrival times deviate from the data. These deviations are found to be systematic and of comparable order of magnitude over a set of 48 such column experiments. Also shown in Figure 14 is a fit to the data using a BTC computed with the η function described above. It can be seen that the solution of (33) with a memory function defined by the η function captures all of the "anomalous" features (in contrast to the ADE model) and also better quantifies the position of the mean arrival.

[96] On the other hand, the Fickian evolution of a BTC for a long column (L = 101 cm) filled with coarse sand was captured correctly using the ADE model [*Cortis et al.*, 2004a]. From this analysis we conclude that there is a transition between the anomalous behavior observed in the short-column experiments and the Fickian behavior characteristic of the long column (discussed below). The anomalous behavior in this case is not due to a $\psi(t)$ with a power law tail but rather to deviation from a pure exponential form (recalling that the pure exponential form gives rise to Gaussian behavior).

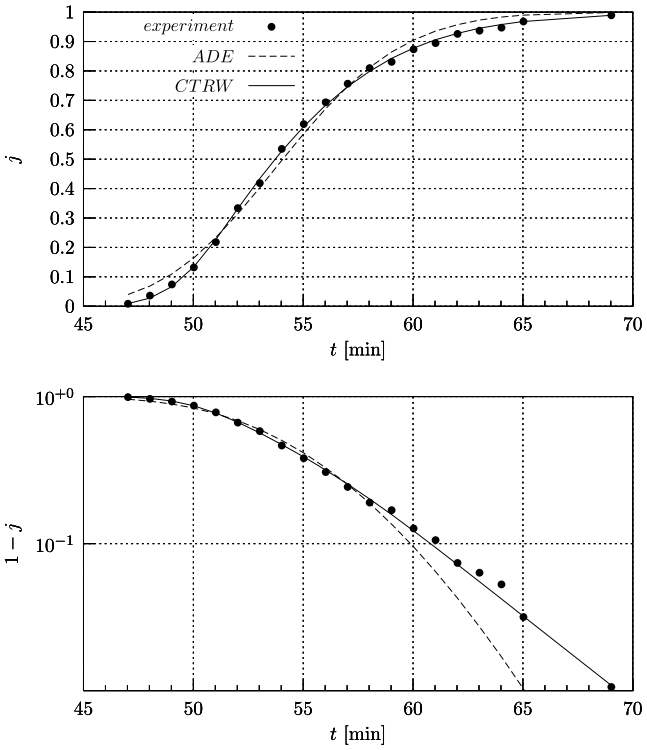

Figure 14. Comparison of measured versus fitted BTCs for a typical short-column experiment, using the η function (21) with η = 0.9212. (top) BTC where here the quantity j represents the normalized, flux-averaged concentration. Dots indicate measured chloride BTC for a fine sand with average grain size diameter of d_m = 0.231 mm. Measured porosity n = 0.3228. Column length L = 19.85 cm. Section area is 4.956×10^{-4} m^2. Volumetric fluid flow $Q = 6.717 \times 10^{-7}$ m^3/min. Dashed line is the best ADE model fit: $\langle v\rangle = 4.917 \times 10^{-3}$ m/min and $D = 3.2363 \times 10^{-6}$ m^2/min. Solid line is the best CTRW fit: $v_\psi = 4.0389 \times 10^{-3}$ m/min and $D_\psi = 1.1601 \times 10^{-6}$ m^2/min. (bottom) Quantity $(1 - j)$ in logarithmic units to emphasize the long time tail. Reprinted with permission from *Cortis et al.* [2004a] (http://dx.doi.org/10.1103/PhysRevE.70.041108). Copyright 2004 by the American Physical Society.

[97] Fitting numerical solutions of (33) to these experimental BTCs, we obtained independently the parameters v_ψ, D_ψ, and η corresponding to the η function. Also we obtained the parameters $\langle v\rangle$ and D for the ADE (i.e., the solution of (33) for $\tilde{M}(u) = 1$). In Figure 15 we plot the value of this "disorder parameter" η against the ratio $\langle v\rangle/v_\psi$ for the set of 48 BTCs. It can be seen that the prediction of (52) is clearly satisfied by the experiments. The fitted values of η range between 0.8 and 1.5. This range of values is characteristic of a weak disorder.

[98] The approximations made in the determination of η are consistent with the values needed to fit the data in a system with small disorder. The robustness of the linear dependence of η in (52) on the ratio of fluid velocity to transport velocity, $\langle v\rangle/v_\psi$, is an a posteriori justification of

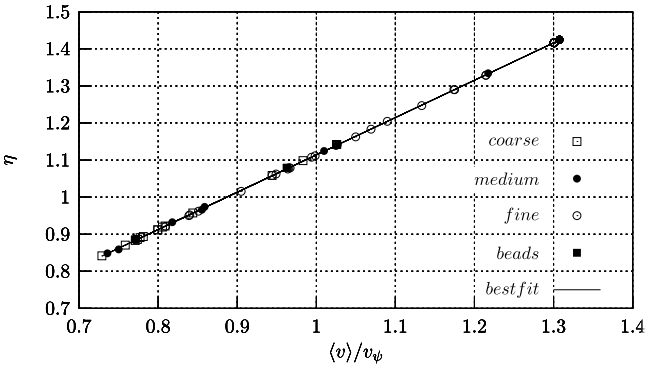

Figure 15. Coefficient η versus the dimensionless velocity $\langle v\rangle/v_{\psi}$ for a set of 48 experimental BTCs on short columns filled with well sorted granular materials of different grain size. Reprinted with permission from *Cortis et al.* [2004a] (http://dx.doi.org/10.1103/PhysRevE.70.041108). Copyright 2004 by the American Physical Society.

the coarse-graining approach and the form of (50) and (51). Moreover, the value of unity obtained for a can be proved theoretically at least for smooth porous media. The importance of these results lies in demonstrating that mild fluctuations from a completely "homogeneous" porous medium have clear effects on the basic observations of transport in these media. Thus the physical picture forming the basis of the classical ADE leads to an incomplete description of transport phenomena even in "homogeneous" media, whereas a CTRW formulation can quantify this kind of transport. The nature of the "non-ADE" behavior analyzed here is different from the non-Fickian behavior discussed in this paper. The effective ψ (t) we use here ($\eta \sim 1$) deviates from a pure exponential ($\propto \exp(-Wt)$) in an incremental way, with no power law tail. Yet as we have discovered, this difference from that required for the use of an ADE clearly has measurable consequences.

[99] The key issue, to reiterate, is to recognize that short-range "geometrical" heterogeneity in a porous medium leads also to short-range spatial variability in the microscopic velocity field $v(y)$ of the liquid that carries the tracer. These latter effects are eventually "averaged out" in, e.g., sufficiently long columns L of granular matter: $L \ggg d_m$, where d_m is a characteristic diameter of the individual grains. However, in a large range where $L \gg d_m$ these fluctuations can produce the measurable, systematic deviations from the BTC derived from the ADE. The volumetric fluxes used in the two tracer experiments were comparable ($Q_1 = 6.717 \times 10^{-7}$ m^3/min and $Q_2 = 7.8760 \times 10^{-7}$ m^3/min for the short (L = 20 cm) and long (L = 101 cm) columns, respectively), such that the residence time of the tracer in the long column ($t_2 \sim 265$ min) was thus roughly 5 times larger than that for the short column ($t_1 \sim 53.7$ min). The ratio of the total length of the column to the diameter of the characteristic disorder (the diameter of the grains) is equal to L/d_m = 870 for the short-column (fine sand) experiment, and it is of the same order of magnitude for the long column (coarse sand) experiment, L/d_m = 909. Experiments on the long column, which showed normal transport behavior, and on the short column, which showed anomalous behavior, clearly demonstrate that the relative residence time is ultimately what governs the transition from anomalous to Fickian behavior. Considerations only of the length-scale separation are not sufficient to discriminate between anomalous and normal transport behavior.

3.6. Fractal Chemistry in Catchment Basins

[100] In a very different type of field study, *Kirchner et al.* [2000] measured the chloride tracer concentration $c_R(t)$ time series in the rainfall over a catchment area in Plynlimon, Wales, and compared it to the time series of the chloride tracer concentration $c_s(t)$ in the catchment Hafren stream. They relate the concentrations through the convolution integral

$$c_s(t) = \int_0^{\infty} h(t')c_R(t-t')dt', \tag{53}$$

where the effective traveltime distribution $h(t)$ governs the lag time between injection of the tracer through rainfall and outflow to the stream. *Kirchner et al.* [2000] performed a spectral analysis of the time series data and showed that the chloride concentrations in rainfall have a white noise spectrum, while in streamflow the spectrum exhibits a fractal $1/f$ scaling, f denoting frequency. The dynamics of the passive tracer transport in the catchment $h(t)$ were extracted from the spectral analysis to conclude that

$$h(t) \sim t^{-m}, \tag{54}$$

where $m \approx 0.5$. Remarkably, *Kirchner et al.* [2000] established this power law behavior over 3 decades of time from 0.01 to 10 years; they also refer to similar scaling with m between 0.4 and 0.65 found in Scandinavian and North American field sites, which indicates a certain ubiquity to fractal, scale-free forms. We account for this behavior with the CTRW framework [*Scher et al.*, 2002b] because the transport is dominated by subsurface flow of the catchment basin, which can be modeled as a heterogeneous porous medium and/or a random fracture network.

[101] The first step to understanding transport in the catchment is the clarification of the meaning of (53). The $h(t)$ is the effective response to a pulse of rain falling on the entire area of the catchment. Every point of this area is a source of chloride, and the stream is a line sink for the chloride. We simplify the area to be a rectangle of width 2λ about this stream sink. The sink or absorbing boundary ensures the correct "counting rate" of chloride at the arriving point; that is, the first-passage time distribution $F(\mathbf{l}, t)$ given by (11) is the appropriate traveltime distribution from a pulse source at the origin to the point $\mathbf{l}$. In terms of this intrinsic distribution,

$$c_s(\mathbf{l}_s, t) = \int_0^{\infty} \sum_{\mathbf{l}\in\Omega} F(\mathbf{l}_s - \mathbf{l}, t')c_R(\mathbf{l}, t-t')dt', \tag{55}$$

where $c_s(\mathbf{l}_s, t)$ is the chloride concentration at the stream position $\mathbf{l}_s$ at time t, Ω is the size of the catchment, and $c_R(t; l)$ is the rain input at a position $\mathbf{l}$ in Ω.

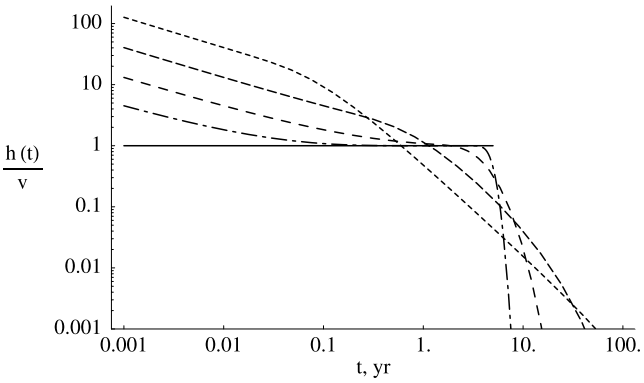

Figure 16. A log-log plot of $h(t)/v$ versus t in years for v = 100 m/yr and the width of the basin λ = 500 m. The curves correspond to Pe of 0.1 (short dashed), 1.0 (medium dashed), 10.0 (long dashed), and 100.0 (dashed-dotted); solid line indicates infinity. The curves for $t < 0.01$ scale as $h(t)/v \sim t^{-1/2}$; this constant slope changes over the range of the experiment, 0.1–10 years, even for very low Pe.

[102] In the catchment being studied, the magnitude of the stream velocity is much faster than the groundwater flow, $v_s \gg v$. Hence, at a sample position downstream one has an "instant" integration of all of the tracer that hits the full length of the stream. Thus it suffices to consider the total tracer flux into the stream integrated over the directions perpendicular to the groundwater flow, which defines the first-passage time distribution at the absorbing boundary. We can consider the sampling positions $\{\mathbf{l_s}\}$ of c_s to be a small region compared to Ω, and hence $c_s(\mathbf{l_s}, t) \approx c_s(t)$. To calculate the summation over $\mathbf{l}$, we assume that the rainfall is distributed uniformly in Ω, $c_R(\mathbf{l}, t) \approx c_R(t)$. Hence we recover (53) now with

$$h(t) \equiv \sum_{\mathbf{l}\in\Omega} F(\mathbf{l}, t). \tag{56}$$

The distribution $F(\mathbf{l}, t)$ must be integrable in time, and for a finite Ω, $h(t)$ is also integrable. The basis of comparison for various transport approaches is the computation in (56).

[103] We compare two approaches to the determination of $F(\mathbf{l}, t)$ in (56) based on the solution of the ADE and CTRW equations, respectively. The simple geometry we have chosen, with periodic BCs on the sides perpendicular to the stream, allows the grouping of catchment chloride sources into line sources parallel to the stream. Hence we consider the one-dimensional form of $F(l_x, t)$, with v the flow component perpendicular to the stream.

[104] The ADE solution is

$$F(l_x, t) = \frac{l_x}{\sqrt{4\pi D}t^{\frac{3}{2}}} \exp\left[\frac{-(vt - l_x)^2}{4Dt}\right], \tag{57}$$

which yields for (56), in one dimension,

$$h(t)/v \propto \frac{1}{2}\left\{\operatorname{erf}(z) + \operatorname{erf}\left(\frac{Pe}{4z} - z\right) + \left[\exp(-z^2) - \exp\left(-\left(\frac{Pe}{4z} - z\right)^2\right)\right]/\pi^{\frac{1}{2}}z\right\}$$

with

$$z \equiv \left(\frac{vt}{4\lambda}\right)^{\frac{1}{2}} Pe^{\frac{1}{2}},$$

where here the Peclet number $Pe \equiv \lambda/\alpha$, the dispersivity $\alpha \equiv D/v$, D is the dispersion coefficient, and erf(z) is the error function.

[105] In Figure 16, $h(t)/v$ is plotted versus t for a range of Pe and fixed values of v = 100 m/yr and λ = 0.5 km. For $t \leq$ 0.01 year, $h(t)/v \sim t^{-\frac{1}{2}}$ for all Pe; however, in the observational span of $0.01 \leq t \leq 10$ years, $h(t)/v$ does not follow this time dependence and is not in accord with the experimental data. For a low value of v = 10 m/yr the Pe = 1.0 curve (not shown here) can just accommodate the data; however, Pe = 1.0 requires an extremely large dispersivity α = 0.5 km. Even with these limiting values the ADE solution can only match the m = 0.5 case (recall (54)), while other catchment basin data quoted above cover a range $0.4 \leq m \leq 0.65$. (Note that for similar limiting parameters, (57), which is the ADE solution for the BTC for a pulse source [e.g., *Kreft and Zuber*, 1978], can exhibit a $t^{-3/2}$ tail.)

[106] It is more expedient in the CTRW computation to work with the $\mathcal{L}$ of $F(x, t)$, which can be found in (47) (based on (44)) with $L \rightarrow l$. The $\mathcal{L}$ of (56) is

$$\tilde{h}(u) \propto \int_0^{\lambda} \exp\left(-c_\beta l\mu^\beta/\bar{l}\right) dl = \bar{l}'\mu^{-\beta}\left[1 - \exp\left(-\lambda\mu^\beta/\bar{l}'\right)\right], \tag{58}$$

where $\bar{l}$ is the mean step distance ($\bar{l}' \equiv \bar{l}/c_\beta$). The expression for $\tilde{h}(u)$ in (58) has been thoroughly studied in another context [*Scher and Montroll*, 1975]; the main features are

$$h(\tau) \sim \begin{cases} \tau^{\beta-1}, & \tau < \tau^* \\ \tau^{-\beta-1}, & \tau > \tau^* \end{cases} \tag{59}$$

with $\tau \equiv t/\xi_o\bar{l}'\varepsilon$, where ξ_o is a characteristic value of the ξ distribution (see (41)), which we take as $\xi_o^{-1} = v$, and ε is a constant. The exponent for $\tau > \tau^*$ ensures that $h(\tau)$ is integrable. On a log-log plot, $h(\tau)$ displays two constant slopes, $\beta - 1$ and $-1 - \beta$, with a turnover range between them. The center time of this range τ^* can be estimated as the time for the argument of the exponent in (58) to be $\sim O(1)$ (using $\mu \sim 1/\tau$)

$$\tau^* \sim \left(\frac{1-\beta}{\beta}\right)^{\frac{1-\beta}{\beta}} \left(\beta\lambda/\bar{l}'\right)^{\frac{1}{\beta}}. \tag{60}$$

(The β factors derive from a more detailed analysis [*Scher and Montroll*, 1975, Appendix C].)

[107] For $\beta = \frac{1}{2}$, in (58) one can determine the inverse $\mathcal{L}$ and obtain $h(\tau) \sim \bar{l}'(\pi\tau)^{-\frac{1}{2}}[1 - \exp(-\lambda^2/4\bar{l}'^2\tau)]$ shown in Figure 17. For $v \sim 100$ m/yr, $\lambda \sim \frac{1}{2}$ km, $\bar{l}' \sim 30$ m, and $\varepsilon \sim \frac{1}{2}$ one has (from (60) and $\tau^*\bar{l}'\varepsilon/v \equiv t^*$) that $t^* \sim 10$ years, which is a reasonable timescale for the change from $h(t) \sim t^{\beta-1}$; we use $h(t)$ and $h(\tau)$ interchangeably. Hence the scaling result (59) agrees with the *Kirchner et al.* [2000] data in (54) over the measurement time range (>3 decades

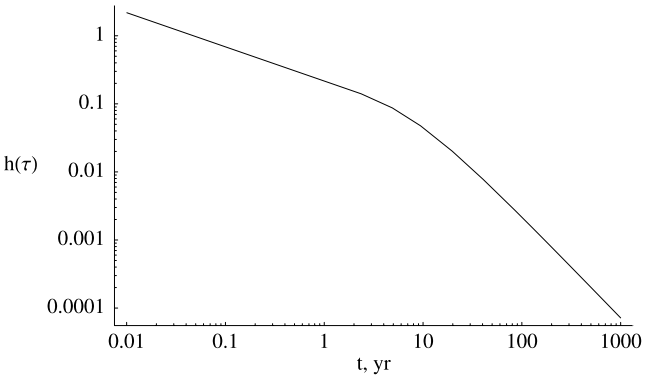

Figure 17. A log-log plot of $h(\tau)$ versus t, with $t \equiv \tau\bar{l}\varepsilon/v$.

for $t \leq 10$ years) with $\beta = 1 - m$. The t "cutoff" for $h(t)$ in (59) is algebraic, $t^{-\beta-1}$, not exponential as given by *Kirchner et al.* [2000]. This behavior of $h(t)$ is indicative of extremely long chemical retention times in catchments. The turnover to the $t^{-\beta-1}$ dependence (in (59)) is a prediction of the CTRW theory and has not yet been observed.

[108] Another prediction of the CTRW theory, to independently test the model (as discussed by *Stark and Stieglitz* [2000]), is the response at the stream to a tracer spill at a site in the catchment. Again, we must calculate the one-dimensional first-passage time distribution to the stream but from a fixed site distance l_x [*Dentz et al.*, 2004]. In this calculation we use a modification of the power law tail (45) for $\psi(t)$. We truncate the tail at a time t_2 because we assume that at such long chemical retention times the transport could evolve to normal behavior. The use of this "truncated" power law $\psi(t)$ is discussed thoroughly in section 3.7.

[109] Figure 18 shows the calculated arrival time distribution (flux) at the stream for $t_2 = 100$ years and tracers originating from three different points along the l_x axis at $l_x = 25$, 100, and 250 m. For this scenario the distance from the absorbing boundary (stream) must be smaller than 80 m in order to observe a pronounced power law behavior. For $l_x = 25$ m the maximum of the arrival time distribution is at about $t = 0.1$ year. Power law behavior is observed for times large compared to 0.1 year and small compared to the cutoff time $t_2 = 100$ years. For $l_x = 100$ m the maximum of the arrival time distribution is shifted to 1 year. One observes tailing of the arrival time distribution for times between 1 and $t_2 = 100$ years, but a distinct power law tail does not develop. For $l_x = 250$ m the maximum is located at about 10 years. The arrival time distribution shows a weak tailing between 10 and 100 years. As expected, the three curves decrease quickly for times larger than the cutoff timescale of 100 years.

[110] In these field observations of tracer transport in catchments, *Kirchner et al.* [2000] demonstrated power law behavior over at least 3 decades of time. A CTRW-based model [*Scher et al.*, 2002b] has accounted for this behavior over such a large time regime and for a range of values of m. Observing the tracer flux into the stream for a site injection at $l_x \sim 25$ m (to detect the power law tail) would be added confirmation of the model proposed by *Scher et al.* [2002b], as discussed by *Stark and Stieglitz* [2000].

3.7. Truncated Power Law

[111] In section 2.5 we derived a $\psi(t)$, (21), that exhibited a number of the general features of the range of transition rates in a disordered system (Figure 3). In section 3.5 we used the large ζ (quasi-exponential) evaluation of this $\psi(t)$ to characterize a nearly homogeneous system. For most of the other applications in section 3 we used a $\psi(t)$ with an asymptotic power tail, (45). The BTC's computed with the input of a power law for $\psi(t)$ in many cases encompass the full set of measurements. The large "power law" region is thus a practical one. However, notwithstanding this accounting for observations, we recognize that (44) has limits: At large enough time in any real physical system the transport becomes normal if the scale of observation is larger than the largest heterogeneity scale. For example, the lower ζ curves in Figure 3 exhibit a power law tail behavior for a range of large values of τ, but with increasing τ, there is a departure from this behavior, and $\psi(t)$ eventually decays faster than τ^{-3}. In other words, there is a cutoff timescale, and transport in the long time limit evolves to Fickian (macrodispersion regime). We stress that the choice of $\psi(t)$ for any given application must of course be motivated by the physical nature of the system under consideration.

[112] A simple version of $\psi(t)$ that takes account of this cutoff behavior is a truncated power law distribution, which enables detailed investigation of the transition from anomalous to normal transport [*Dentz et al.*, 2004]:

$$\psi(t) = \left[t_1\,\tau_2^{-\beta}\,\exp\left(\tau_2^{-1}\right)\Gamma\left(-\beta,\tau_2^{-1}\right)\right]^{-1}\frac{\exp(-t/t_2)}{(1+t/t_1)^{1+\beta}}, \qquad (61)$$

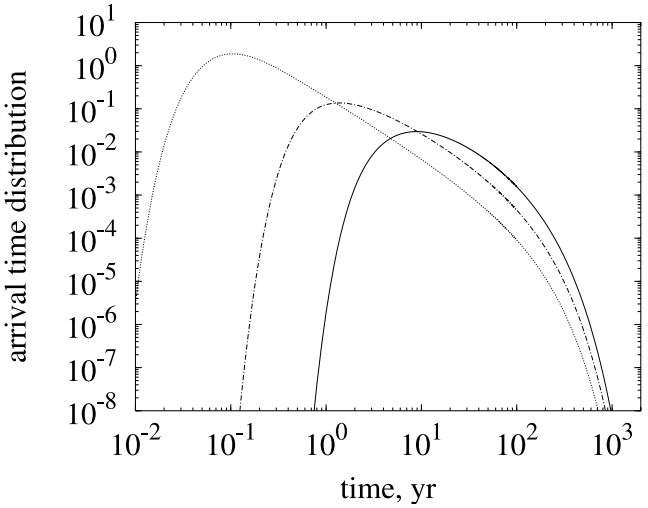

Figure 18. Arrival time distributions at the absorbing boundary ($l_x = 0$) for $t_2 = 100$ years and $l_x = 25$ m (dotted curve), $l_x = 100$ m (dash-dotted curve) and $l_x = 250$ m (solid curve). Reprinted from *Dentz et al.* [2004]. Copyright 2004, with permission from Elsevier.

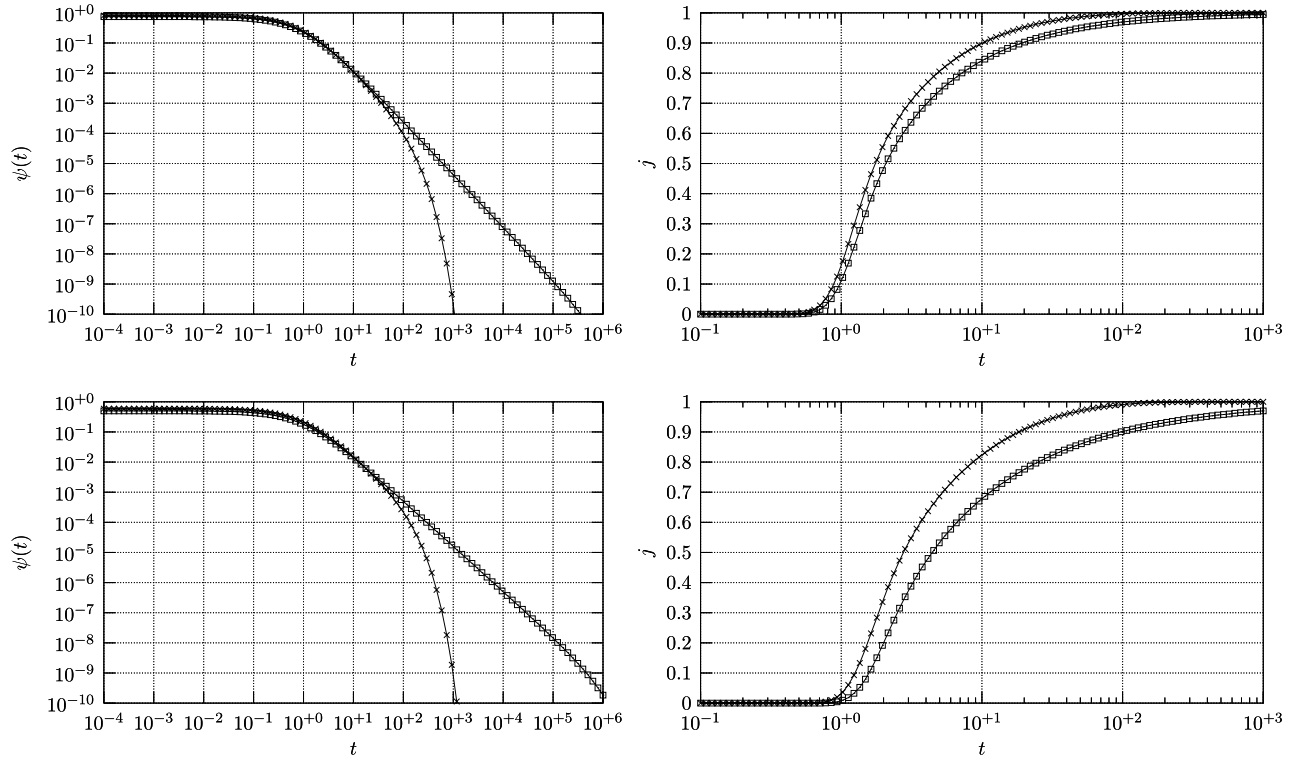

Figure 19. Behavior of the truncated power law (61) $\psi(t)$ function and the corresponding BTCs (the quantity j represents the normalized, flux-averaged concentration). The behavior of $\psi(t)$ and the BTCs is shown for (top) $\beta = 0.75$ and (bottom) $\beta = 0.5$. The function (61) is plotted for two different values of the truncation time τ_2, with $t_1 = 1$ (in an appropriate time unit) ($\tau_2 = 10^2$ (crosses) and $\tau_2 = 10^6$ (squares)). The BTCs are calculated for one-dimensional flow over a unit domain, with a free flow BC at the outlet. After *Cortis et al.* [2004b].

where $\tau_2 \equiv t_2/t_1$ and $\Gamma(a, x)$ is the incomplete Gamma function [*Abramowitz and Stegun*, 1970]. The $\mathcal{L}$ of (61) is given by

$$\tilde{\psi}(u) = (1 + \tau_2 u t_1)^{\beta} \exp(t_1 u) \cdot \Gamma\left(-\beta, \tau_2^{-1} + t_1 u\right)/\Gamma\left(-\beta, \tau_2^{-1}\right). \quad (62)$$

The $\tilde{M}(u)$ defined in section 2.6 is determined by the substitution of (62) into (29), with $\bar{t} = t_1$. The time t_1 represents the approximate median transition time and sets the lower limit from which power law behavior begins. Figure 19 shows the behavior of $\psi(t)$ for different values of β and for different truncation times.

[113] It is highly instructive to investigate the complete transport behavior of a particle for the truncated power law distribution that is characterized by the exponent β and by the two timescales t_1 and t_2. For $t_1 \ll t \ll t_2$, $\psi(t) \propto (t/t_1)^{-1-\beta}$ as in (45). In this time regime the transport behavior is anomalous for $0 < \beta < 2$ [*Shlesinger*, 1974]. The presence of a cutoff time t_2 ensures that for $t > t_2$ the transport evolves into a normal one. Hence anomalous transport in any system resides in the interplay between the duration of the observation and the extent of disorder, which is reflected in the range of t up to the cutoff t_2.

[114] The exponent β is a very useful means to characterize the latter range. However, we have introduced the truncated power form in (61) and (62) to emphasize the important point that all physically meaningful forms of $\psi(t)$ must have a large dimensionless τ ($\equiv t/t_1$) behavior that causes the transport to evolve inevitably to a normal one. The truncated power form allows for a sharp transition $\tau > \tau_2$ to the limit of normal transport, whereas the form in (21) exhibits a more gradual transition.

3.7.1. Random Walk Simulations

[115] To assess quantitatively use of the CTRW in pde form (section 2.6) and numerical inversion of Laplace transforms (section 2.4), we compare the results to random walk simulations carried out using (61) and a $p(\mathbf{s})$ with finite moments [*Dentz et al.*, 2004]; see also Appendix A. It suffices, without loss of generality (because of the central limit theorem), to choose for $p(\mathbf{s})$ a Gaussian distribution,

$$p(\mathbf{s}) = \frac{1}{\sqrt{2\pi\sigma_L}} \exp\left[-\frac{(s_1 - \overline{s_1})^2}{2\sigma_L}\right] \prod_{i=2}^{d} \frac{1}{\sqrt{2\pi\sigma_T}} \exp\left[-\frac{s_i^2}{2\sigma_T}\right], \quad (63)$$

where σ_L and σ_T are the longitudinal and transverse components of the variance $\boldsymbol{\sigma}$, respectively, which we assume to be diagonal. Here we choose $\sigma_L = \sigma_T$. Lengths are measured in units of $\overline{s_1}$.

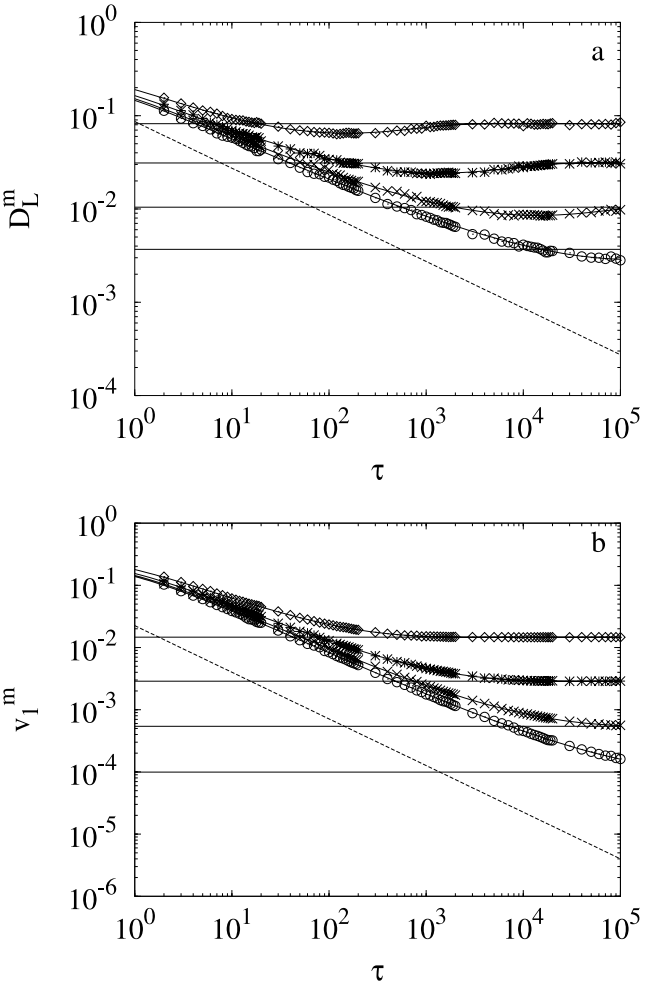

Figure 20. Simulated time behavior of the (a) longitudinal dispersion coefficient and (b) center of mass velocity and transverse dispersion coefficients for $\beta = 0.25$, $\tau_2 = 10^3$ (diamonds), 10^4 (asterisks), 10^5 (crosses), and 10^6 (circles). The horizontal lines are the corresponding macrodispersion regime values. The solid curves result from the inverse $\mathcal{L}$ of (67) and (68). The dashed curves in Figures 20a and 20b represent asymptotic approximations [*Shlesinger*, 1974; *Dentz and Berkowitz*, 2003]. Reprinted from *Dentz et al.* [2004]. Copyright 2004, with permission from Elsevier.

[116] The equations of motion for the random walk simulations are given by

$$\mathbf{s}^{(N+1)} = \mathbf{s}^{(N)} + \boldsymbol{\xi}^{(N)} \tag{64a}$$

$$t^{(N+1)} = t^{(N)} + \theta^{(N)}, \tag{64b}$$

where $(\mathbf{s}^{(N)}, t^{(N)})$ denotes the location of a particle in space and time after N steps. The spatial and temporal random increments $\boldsymbol{\xi}^{(N)}$ and $\theta^{(N)}$ are distributed according to $p(\mathbf{s})$ and $\psi(t)$, i.e., as given in (61) and (63). The simulation results are shown in Figures 20–26 and are compared to calculations of the analytic solutions of the CTRW discussed in sections 3.7.2 and 3.7.3.

3.7.2. Center of Mass Velocity and Dispersion Coefficients

[117] The center of mass velocity, $v_i^m(t)$, and the dispersion coefficients, $D_{ij}^m(t)$, are derived from the first

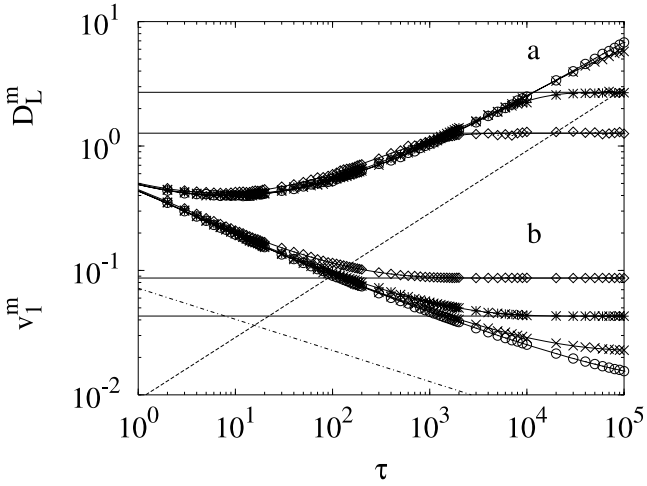

Figure 21. Simulated time behavior of the longitudinal dispersion coefficient (labeled a) and the center of mass velocity and transverse dispersion coefficients for $\beta = 0.75$, $\tau_2 = 10^3$ (diamonds), 10^4 (asterisks), 10^5 (crosses), and 10^6 (circles) (labeled b). The horizontal lines are the corresponding macrodispersion regime values. The solid curves result from the inverse $\mathcal{L}$ of (67) and (68). The dashed curves represent the asymptotic approximations [*Shlesinger*, 1974; *Margolin and Berkowitz*, 2002; *Dentz and Berkowitz*, 2003]. Reprinted from *Dentz et al.* [2004]. Copyright 2004, with permission from Elsevier.

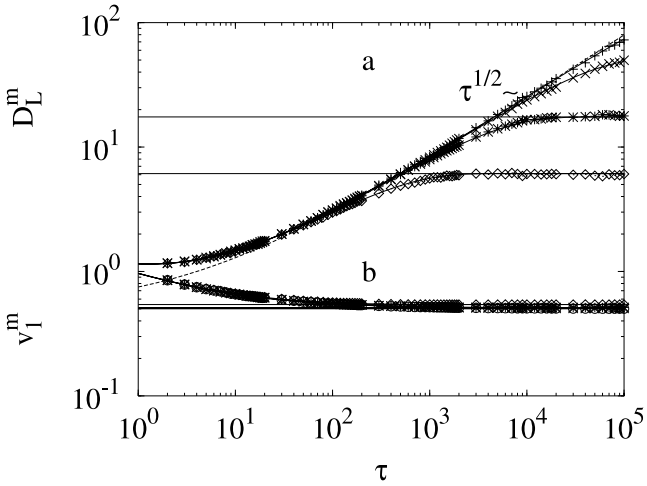

Figure 22. Simulated time behavior of the longitudinal dispersion coefficient (labeled a) and the center of mass velocity and transverse dispersion coefficients for $\beta = 1.5$, $\tau_2 = 10^3$ (diamonds), 10^4 (asterisks), 10^5 (crosses), and 10^6 (plus signs) (labeled b). The horizontal lines are the corresponding macrodispersion regime values. The solid curves result from the inverse $\mathcal{L}$ of (67) and (68). The dashed curve represents the asymptotic approximation [*Shlesinger*, 1974; *Margolin and Berkowitz*, 2002; *Dentz and Berkowitz*, 2003]. Reprinted from *Dentz et al.* [2004]. Copyright 2004, with permission from Elsevier.

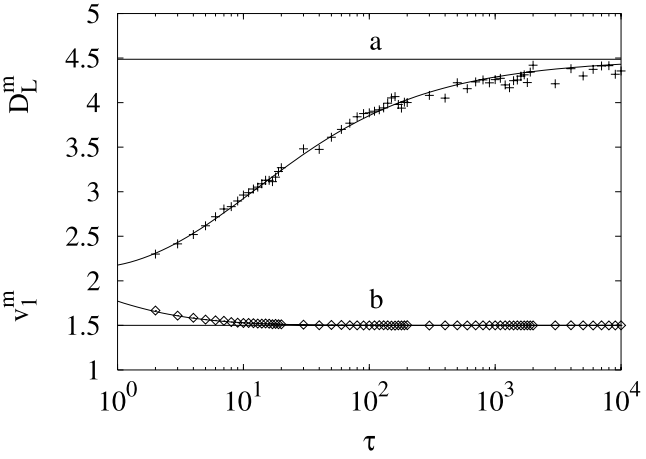

Figure 23. Simulated time behavior (plus signs and diamonds) of the longitudinal dispersion coefficient (labeled a) and the center of mass velocity and transverse dispersion coefficients (labeled b) for $\beta = 2.5$ and $\tau_2 = 10^6$. The horizontal lines are the corresponding macrodispersion regime values. The solid curves result from the inverse $\mathcal{L}$ of (67) and (68) for $\tau_2 = 10^6$. Reprinted from *Dentz et al.* [2004]. Copyright 2004, with permission from Elsevier.

$m_i^{(1)}(t)$ and second spatial moment $m_{ij}^{(2)}(t)$ of $c(\mathbf{s}, t)$. Without loss of generality we choose the velocity to be aligned with the one direction of the coordinate system, with v the magnitude of the velocity. In dimensionless form,

$$v_i^{\mathrm{m}}(t) = v^{-1} \frac{\mathrm{d}}{\mathrm{d}t} m_i^{(1)}(t) \tag{65}$$

$$D_{ij}^{\mathrm{m}}(t) = D_{ij}^{-1} \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \left[m_{ij}^{(2)}(t) - m_i^{(1)}(t)\, m_j^{(1)}(t) \right] \tag{66}$$

for $v \neq 0$ and $D_{ij} \neq 0$, where the notation is the same as used in section 2.6, before (37), and the superscript m denotes

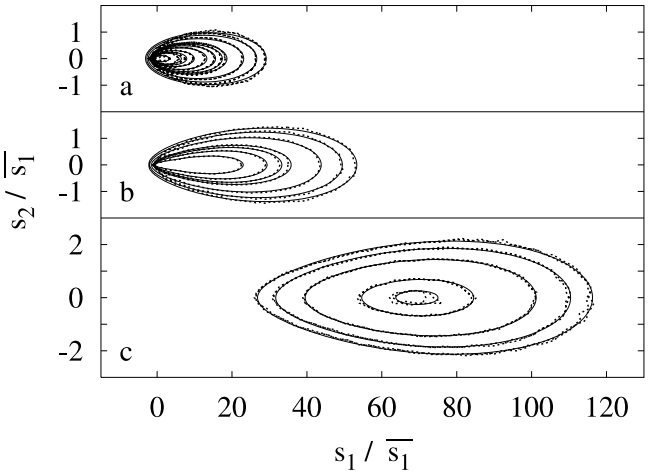

Figure 24. Resident concentrations $c(\mathbf{s}, t)$ at $\tau = 10^3$ for $\beta = 0.25$ in two spatial dimensions: (a) $\tau_2 = 10^5$, (b) $\tau_2 = 10^3$, and (c) $\tau_2 = 10^2$. The solid curves represent the behavior resulting from the inverse $\mathcal{L}$ of (38). Dots represent numerical simulations using the random walk given in (64a) and (64b). Reprinted from *Dentz et al.* [2004]. Copyright 2004, with permission from Elsevier.

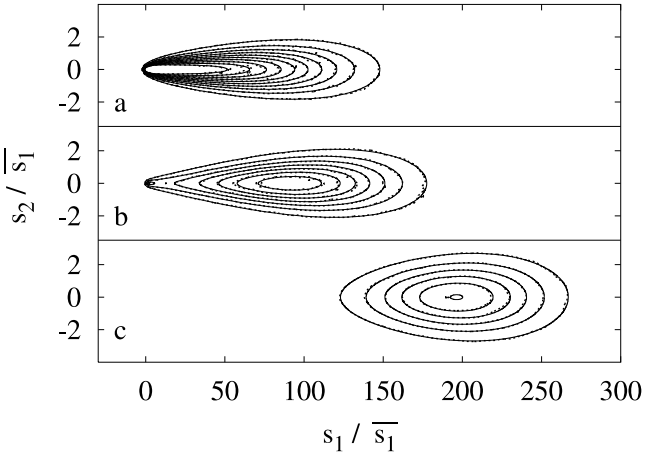

Figure 25. Resident concentrations $c(\mathbf{s}, t)$ at $\tau = 10^3$ for $\beta = 0.75$ in two spatial dimensions: (a) $\tau_2 = 10^5$, (b) $\tau_2 = 10^3$, and (c) $\tau_2 = 10^2$). The solid curves represent the behavior resulting from the inverse $\mathcal{L}$ of (38). Dots represent numerical simulations using the random walk given in (64a) and (64b). Reprinted from *Dentz et al.* [2004]. Copyright 2004, with permission from Elsevier.

"macroscopic" quantities. Using formula (36) and the definition of the memory function (29), we derive in Appendix D of *Dentz et al.* [2004] the $\mathcal{L}$ of the first and second moments:

$$\mathcal{L}\left[m_i^{(1)}(t) \right] = \delta_{i1}\, v\, u^{-2} \tilde{M}(u) \tag{67}$$

$$\mathcal{L}\left[m_{ii}^{(2)}(t) \right] = 2\, D_{ii}\, u^{-2} \tilde{M}(u) + \delta_{i1}\, 2\, v^2\, u^{-3} \left[\tilde{M}(u) \right]^2. \tag{68}$$

The time behavior of the center of mass velocity and the dispersion coefficients (65) and (66) are calculated by $\mathcal{L}^{-1}$ of (67) and (68) using the explicit expression (62) for the $\mathcal{L}$ of the transition time distribution.

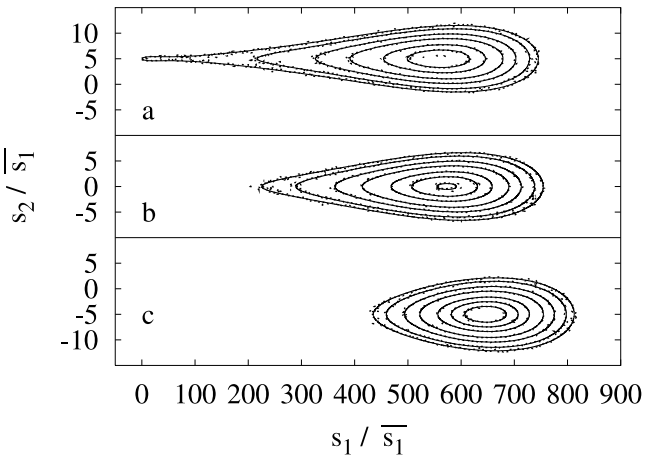

Figure 26. Resident concentrations $c(\mathbf{s}, t)$ at $\tau = 10^3$ for $\beta = 1.5$ in two spatial dimensions: (a) $\tau_2 = 10^5$, (b) $\tau_2 = 10^3$, and (c) $\tau_2 = 10^2$. The solid curves represent the behavior resulting from the inverse $\mathcal{L}$ of (38). Dots represent numerical simulations using the random walk given in (64a) and (64b). Reprinted from *Dentz et al.* [2004]. Copyright 2004, with permission from Elsevier.

[118] The moments (65) and (66) provide valuable information about the spatial shape of the solute plume. However, we stress that they cannot be used directly in a local time ADE. The timescale on which (66) approaches a constant is an indicator of when the transport behavior becomes Gaussian (normal).

[119] The off-diagonal coefficients of the second moment vanish for symmetry reasons because the ''local'' dispersion tensor **D** is assumed to be diagonal. The $\mathbf{D}^{\mathrm{m}}(t)$ is diagonal and, as can be seen from (67) and (68), the transverse elements are equal, $D_{ii}^{\mathrm{m}} = D_{jj}^{\mathrm{m}}(t)$, $i, j > 1$. Thus we employ in the following the notation $D_{11}^{\mathrm{m}}(t) = D_L^{\mathrm{m}}(t)$ and $D_{ii}^{\mathrm{m}}(t) = D_T^{\mathrm{m}}(t)$, $i > 1$. From (65)–(68) we conclude that the time behavior of the (dimensionless) transverse dispersion coefficients is identical to the behavior of the (dimensionless) center of mass velocity:

$$D_T^{\mathrm{m}}(t) = v_1^{\mathrm{m}}(t). \tag{69}$$

[120] In the limit $t \to \infty$ the transport behavior is Fickian [*Dentz and Berkowitz*, 2003]. In this case the solute transport is characterized completely by the (constant) center of mass velocity and dispersion coefficients, which are given by *Shlesinger* [1974] and *Dentz and Berkowitz* [2003]

$$v_i^{\mathrm{m}} = \delta_{i1}\,\frac{t_1}{\hat{t}}, \tag{70a}$$

$$D_{ii}^{\mathrm{m}} = \frac{t_1}{\hat{t}} + \delta_{i1}\,\frac{\overline{s_i}^2}{\overline{s_i^2}}\,\frac{t_1\left(\widehat{t^2} - 2\hat{t}^2\right)}{\hat{t}^3}, \tag{70b}$$

$$\frac{D_L^{\mathrm{m}}(t)}{D_T^{\mathrm{m}}(t)} = 1 + \frac{\overline{s_i}^2}{\overline{s_i^2}}\left(\frac{\widehat{t^2}}{\hat{t}^2} - 2\right). \tag{70c}$$

The mean time $\hat{t}$ and mean squared times for the truncated power law distribution (61) are given in Appendix C of *Dentz et al.* [2004] and are explicit functions of τ_2 and β.

[121] The numerical results are presented in two complementary ways. In section 3.7.3 we exhibit two-dimensional contour plots of the resident concentration (solute plumes) at a fixed time and a range of β, τ_2. Here, to show the full time evolution of these solute plumes, we plot the center of mass velocity and dispersion coefficients $v_1^{\mathrm{m}}(t)$ (and hence $D_T^{\mathrm{m}}(t)$, see (69)) and $D_L^{\mathrm{m}}(t)$, respectively, versus τ and a range of β, τ_2 in Figures 20–23.

[122] The horizontal lines in Figures 20–23 are given by (70a) and (70b) and are the limiting values of all the curves. In Figures 20–23, there is excellent agreement between the simulations and the results of the numerical $\mathcal{L}^{-1}$ of (67) and (68) (based on solutions of the pde form of CTRW). Figures 20–23 are a gallery of behavior each with a value of β representing a distinct range. The τ regions of time dependence of $v_1^{\mathrm{m}}(t)$ and/or $D_L^{\mathrm{m}}(t)$ correspond to anomalous transport, and the regions where both have constant values are normal transport. In Figures 20 and 21, $0 < \beta < 1$, and in the time regime $1 \ll \tau \ll \tau_2$ the time dependence can be estimated by the asymptotic expressions in (46), $v_1^{\mathrm{m}} \propto \tau^{\beta-1}$ (the derivative of $\bar{\ell}(t)$) and $D_L^{\mathrm{m}}(t) \propto \tau^{2\beta-1}$ (the derivative of $\bar{\sigma}^2(t)$). This time dependence strictly only holds for the case (61) for $\tau_2 \to \infty$ i.e., without the influence of a finite τ_2. The dashed lines in Figures 20–23 represent these asymptotic expressions derived from the leading order in the small u expansion of (67) and (68).

[123] In Figure 20, $\beta = 0.25$ and the asymptotic analysis predicts $v_1^{\mathrm{m}} \propto \tau^{-3/4}$ and $D_L^{\mathrm{m}} \propto \tau^{-1/2}$, which describes the time behavior of the observables only qualitatively but accounts for the monotonic decrease for both in the interval $1 \ll \tau \ll \tau_2$. The agreement with the asymptotic approximation improves with increasing τ_2. For $\tau \gg \tau_2$ they approach final (macrodispersion regime) constant values given by (70a) and (70b), respectively, which decrease with increasing τ_2.

[124] Figure 21 shows calculations for $\beta = 0.75$, and the asymptotic analysis predicts $v_1^{\mathrm{m}} \propto \tau^{-1/4}$ and $D_L^{\mathrm{m}} \propto \tau^{1/2}$, which describes the time behavior of the observables only qualitatively but accounts for the monotonic decrease of $v_1^{\mathrm{m}}(t)$ or $D_T^{\mathrm{m}}(t)$ (subdiffusive) and the monotonic increase of $D_L^{\mathrm{m}}(t)$ (superdiffusive) in the interval $1 \ll \tau \ll \tau_2$. On the timescale τ_2 we observe a crossover to the (macrodispersion regime) constant value that is reached for $\tau \gg \tau_2$. This value increases with increasing τ_2 for $D_L^{\mathrm{m}}(t)$ and as in Figure 20 decreases for $D_T^{\mathrm{m}}(t)$. In contrast, for $\beta = 0.25$ the transverse as well as the longitudinal spreading is subdiffusive; here D_L^{m} is superdiffusive and D_T^{m} is subdiffusive. (Note that for the purely diffusive case ($v = 0$) the t dependence of $D = s^2/t$ is $t^{\beta-1}$, which always decreases (subdiffusive) for $\beta < 1$ and increasing t.)

[125] An important consequence of the behavior shown in Figures 20 and 21 is the effect of the time duration (and hence heterogeneity scale) of the anomalous regime on the final shape of the normal plume. The ratio of longitudinal dispersion to transverse dispersion (70) is an explicit function of τ_2 and β through $\widehat{t^2}/\hat{t}^2$ (see Appendix C of *Dentz et al.* [2004]). From Figures 20 and 21 one can discern that the ratio increases as τ_2 increases for $1/2 < \beta < 1$ and for $1 < \beta < 2$. In Figure 20 ($0 < \beta < 1/2$) one can note a small increase with τ_2. While the emphasis in the literature has been on the macrodispersion, here we have shown the close connection between the ratio $D_L^{\mathrm{m}}/D_T^{\mathrm{m}}$ and their relation to observables (e.g., breakthrough curves) at earlier times.

[126] For $1 < \beta < 2$ the center of mass velocity (v_1^{m}) is constant (the timescale τ_2 is not relevant in this case), while the asymptotic approximation yields $D_L^{\mathrm{m}}(t) \propto \tau^{2-\beta}$. For $\beta = 1.5$ in Figure 22, $D_L^{\mathrm{m}} \propto \tau^{1/2}$ in the time regime $1 \ll \tau \ll \tau_2$, and the behavior of the observable is well described by the asymptotic approximation.

[127] For $\beta > 2$ both v_1^{m} and dispersion coefficients are constant to leading order. Figure 23 shows the time behavior of D_L^{m}, v_1^{m} for $\beta = 2.5$ and $\tau_2 = 10^6$. The D_L^{m} increases monotonically. In contrast to the time behavior observed for $0 < \beta < 2$, D_L^{m} approximates its macrodispersion regime value already in the intermediate time regime $1 \ll \tau \ll \tau_2$. The timescale τ_2 is not critical for the time evolution of D_L^{m}.

3.7.3. Resident Concentration

[128] To display the qualitative change in the character of the spatial concentration profiles in the transition from anomalous to normal transport behavior, we fix the observation time τ in the contour plots of $c(\mathbf{s}, t)$, (38),

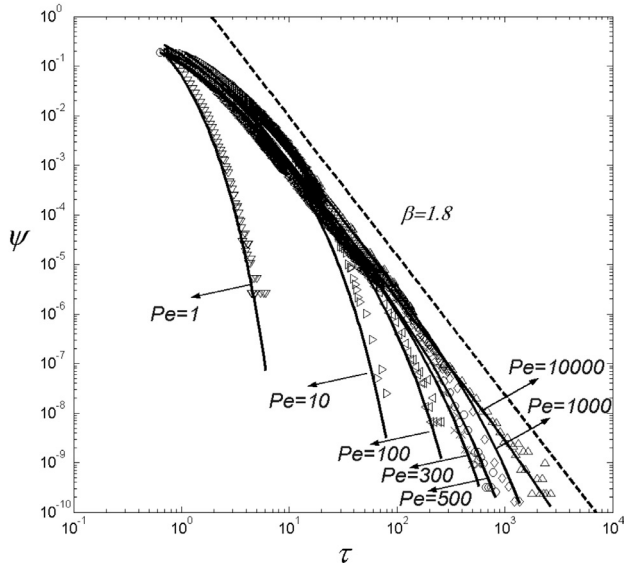

Figure 27. The $\psi(\tau)$ pdf of traveling between two neighboring pores in a dimensionless time $\tau = t/t_1$, where t_1 is the mean traveltime L/v, plotted for different values of *Pe* (denoted by arrows). Also shown (curves) are fits using (61) with a single adjustable parameter $\beta = 1.8$. After *Bijeljic and Blunt* [2006].

in Figures 24 to 26 and vary the cutoff time τ_2. Also shown are the results of the random walk simulation with the same $p(\mathbf{s})$ and $\psi(t)$; they are seen to be in excellent agreement with the contour plots.

[129] The solute evolves from a point-like injection at $\mathbf{s} = \mathbf{0}$ and $t = 0$. Figures 24–26 are a gallery of profiles for $\beta = 0.25$, 0.75, and 1.5 and in each figure $\tau_2 = 10^5$, 10^3, and 10^2 at an observation time of $\tau = 10^3$. Figures 24c–26c indicate normal transport, and the profile approximates the corresponding (macrodispersion regime) Gaussian solution, although with varying degrees of "memory" of the system. In the anomalous limit (Figures 24a–26a), there is an interesting dependence on β. For $\beta < 1$ in Figures 24 and 25 the peak of the profile remains near the point of injection (as shown in the $y = 0$ plane projection in Figure 4) with more localization for $\beta = 0.25$. In Figure 26 for $\beta = 1.5$ the peak moved from the origin (at constant center-of-mass velocity, see Figure 22); however, there is still anomalous behavior with a long backward tail. This tear drop shape has been observed in laboratory experiments as shown in Figures 1 and 10, as well as by *Weisbrod et al.* [2003].

[130] The transition from anomalous to normal transport (Figures 24b–26b) also depends on β, with a very slow change for $\beta = 0.25$. The shape is reminiscent of the one in Figure 25a; in Figure 26b for $\beta = 1.5$ a decline in the asymmetry of the shape is evident.

[131] Depending on the typical transport timescales, here t_1 and t_2, the transport of a solute can display highly non-Gaussian features over a large time regime. Even at extremely large times, where transport is expected to be in the macrodispersion regime, one observes deviations from normal transport patterns. Only in the limit $t \gg t_2$ is Gaussian behavior approximated. Depending on t_2, these times can be so large that they are not of practical interest because the solute is so spread out that the spatial concentration is exceedingly small. Hence, for highly heterogeneous systems the non-Fickian regime is the entire area of significance. Figures 20 to 26 exhibit a rich set of behaviors in this regime.

3.8. Macrodispersion Regime

[132] In those cases where one can observe the macrodispersion regime, we discussed in section 3.7 the crucial dependence of the ratio of D_L^m/D_T^m on the duration of the non-Fickian regime, i.e., the explicit dependence on τ_2. The value of the macrodispersion itself is obviously $D_L^m(\tau_2)$ (see Figures 20–22). Hence it is to be stressed that knowledge of the non-Fickian $D_L^m(\tau)$ and a determination of τ_2 is sufficient to calculate the macrodispersion.

[133] Recent work of *Bijeljic and Blunt* [2006] has demonstrated this relationship convincingly. They obtained numerically the spatial moments (variance σ^2) of a particle transported through flow (average velocity v) and molecular diffusion (D_m) in a pore-scale model [*Bijeljic et al.*, 2004] and then computed the dispersion coefficients according to $D = \sigma^2/2t$. The distribution of times $\psi(t)$ for the particle to move through a throat from one pore to another was fit with (61), with $t_1 = L/v$ being the mean traveltime over the interpore distance L (t_2 is considered below). In Figure 27 this distribution is plotted as a function of τ ($\equiv t/t_1$, as above) for a range of Peclet number, *Pe* ($\equiv vL/D_m$). All the curves are fit with one adjustable parameter, β, using the value $\beta = 1.8$. In section 3.7.2 we stated that for $t_1 \ll t \ll t_2$, $1 < \beta < 2$, $D_L^m(t) \propto \tau^{2-\beta}$ (see Figure 22). In this case, $2 - \beta = 0.2$ and the numerical results in Figure 28 confirm the prediction $D_L^m(t) \propto \tau^{0.2}$. *Bijeljic and Blunt* [2006] also determined $D_T^m(t)$ to be independent of τ in agreement with Figure 22.

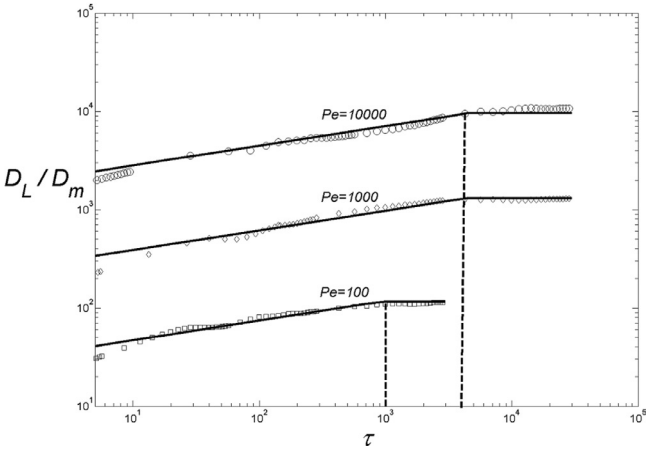

Figure 28. The premacrodispersion regime behavior of D_L computed from pore-scale modeling as a function of dimensionless time τ for different values of Pe (circles). The solid segments indicate the predicted scaling using Figure 27 and CTRW: $D_L/D_m \sim Pe\ \tau^{2-\beta}$ for $t_2/t_1 > \tau \gg 1$ and $1 < \beta < 2$. After *Bijeljic and Blunt* [2006].

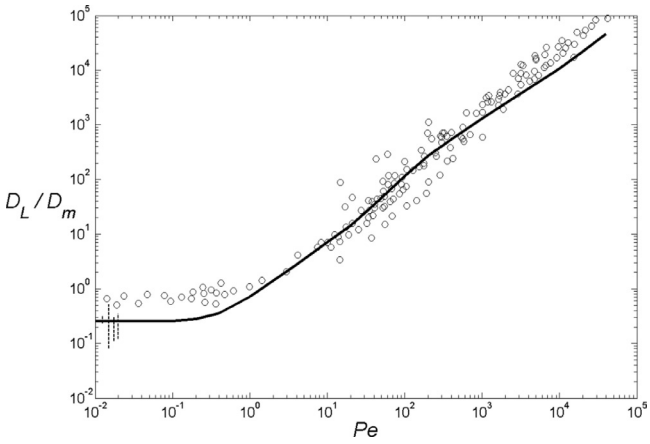

Figure 29. Experimental measurements of longitudinal dispersion coefficient, D_L, as a function of Pe on bead packs, sand packs, and sandstones (circles and dashed vertical lines) compared to predictions using pore-scale modeling (solid curve) [*Bijeljic et al.*, 2004] and CTRW results from Figure 28. For $Pe \gg 1$, there are two scaling regimes: $D_L/D_m \sim Pe^{3-\beta}$ for $Pe^{crit} > Pe \gg 1$ and $D_L/D_m \sim Pe$ for $Pe > Pe^{crit}$, with $Pe^{crit} \approx 400$ and $\beta = 1.8$. After *Bijeljic and Blunt* [2006].

[134] In the notation of section 3.7, $D_L^m(t)$ is dimensionless, and we work in terms of dimensionless time τ. In the work of *Bijeljic and Blunt* [2006], D_L (derived from $\sigma^2/2t$) is made dimensionless by the ratio D_L/D_m. Thus in relation to the current notation we have $D_L/D_m = vLD_L^m(t)/D_m = Pe\,D_L^m(t)$. *Bijeljic and Blunt* [2006] used this relationship to determine the macrodispersion from $D_L(\tau_2)$, as discussed above, where τ_2 is obtained by physical principles. For moderate Pe the largest cutoff interpore transit time is the time to diffuse between pores, $t_2 = L^2/2D_m$, so that $\tau_2 = t_2/t_1 = (L^2/2D_m)\,(v/L) = Pe/2$. For larger Pe this diffusion time is larger than L/v_{min}, where v_{min} is a typically smallest flow velocity. *Bijeljic and Blunt* [2006] found that the crossover between these times occurs when $L^2/2D_m = L/v_{min}$ or $Pe = Pe^{crit} = 2v/v_{min} \sim 400$, where the distribution of throat radii was taken as typical of Berea sandstone. Thus the scaling $D_L/D_m \sim Pe\,\tau_2^{0.2} \propto Pe^{1.2}$; for $Pe > Pe^{crit}$, $D_L/D_m \sim Pe$ because in this range τ_2 is independent of Pe. In Figure 29 this prediction of the scaling of D_L/D_m with Pe yields an excellent account of the experimental measurements of the longitudinal dispersion coefficient on a variety of porous media. The numerical data also confirmed $D_T/D_m \sim Pe$, $Pe \gg 1$. The results of *Bijeljic and Blunt* [2006] take on additional significance in relation to the conceptual framework of ADE-based approaches, which we analyze critically in section 6.

4. NONSTATIONARY DOMAINS

[135] The developments of the preceding sections have considered an ensemble average over the entire domain of the geological medium. Clearly, this is suitable only if the sizes of the significant heterogeneities are much smaller than the domain size. The ensemble-averaged domain is a spatially stationary one. If the distribution of heterogeneities has a large variance, characteristic of most naturally occurring geological formations, large-scale regions of "faster" and "slower" flow and transport are present. Here we consider quantitative treatment of transport in domains that contain heterogeneities that are well defined on the large scale (trends) and unresolved on the small scale (residues) (recall section 2). In these domains an ensemble average is inappropriate, and we must consider, a priori, its nonstationary character. Physical evidence necessitating this approach to treating nonstationary systems is given by, e.g., *Eggleston and Rojstaczer* [1998], *Feehley et al.* [2000], and *Labolle and Fogg* [2001]. These studies emphasize definition of facies structures and their distribution.

4.1. Hybrid, Two-Scale Formulation

[136] We consider a nonstationary medium which has known varying properties at a scale $x \gg \ell$ and unresolved heterogeneities at a scale $y < \ell$ (recall section 2). A schematic showing these different scales is given in Figure 30.

[137] Most modeling efforts are founded on discretization of the domain of interest into "homogeneous" zones with prescribed hydrogeological properties. The division into zones can be conditioned on, e.g., known geological information and geostatistical analyses and can incorporate explicit correlations in hydraulic properties. We shall follow a similar approach but treat transport in each of these "homogeneous" zones with a CTRW formulation as developed in section 2.6. As a result we can integrate resolved and unresolved heterogeneities in a single "hybrid" approach to transport modeling. The term "hybrid" refers to the joining of numerical treatment of large-scale known heterogeneities with a probabilistic treatment of small-scale, unknown residual heterogeneities that play an important role in shaping the overall nature of the transport.

[138] In this hybrid framework the larger-scale known heterogeneities are treated deterministically, and the relevant equation for the mass balance of a nonstationary medium is the classical ME (1) or, more specifically, (27) [*Berkowitz et al.*, 2002]. The unresolved small-scale heterogeneities are treated statistically (as stationary subdomains or local neighborhoods) by the GME (2). Thus the

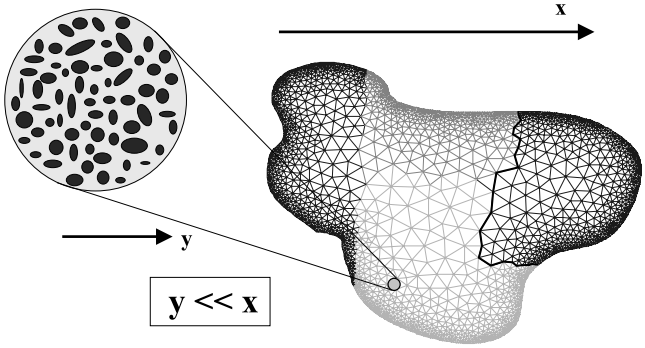

Figure 30. Two-scale system under consideration. The unresolved heterogeneities at the small scale **y** are treated stochastically, whereas the resolved heterogeneities at the large scale **x** are treated deterministically. After *Cortis et al.* [2004b].

developments of sections 2.2 and 2.6, and specifically of (33), are relevant here. Clearly, this resolution scale is somewhat arbitrary, and in practical applications it should be conditioned on the level of detail available from hydrogeological, geophysical measurements and considerations as described by, e.g., *Eggleston and Rojstaczer* [1998], *Weissmann et al.* [1999], and *Lu et al.* [2002].

4.2. Integrating the Small and Large Scales

[139] As detailed by *Cortis et al.* [2004b], we represent the evolution of a contaminant in a two-scale system combining the small-scale GME terms with the large-scale ME into a single equation, hereinafter referred to as the Fokker-Planck with memory equation (FPME). The FPME is a merging of (27) (in Laplace space) and (33). This merging involves the introduction of a position dependence of $\tilde{\psi}(\mathbf{s}, u)$, denoted $\tilde{\psi}(\mathbf{s}, u; \mathbf{x})$; in the decoupled form $p(\mathbf{s}; \mathbf{x})\tilde{\psi}(u; \mathbf{x})$ the $\tilde{\psi}(u; \mathbf{x})$ is considered to be a purely local change in the u dependence and scaling. This formulation gives rise to a space-dependent memory term $\tilde{M}(u; \mathbf{x})$, specific discharge $\mathbf{q}_\psi(\mathbf{x}) = n(\mathbf{x})\mathbf{v}_\psi(\mathbf{x})$ with a space-dependent porosity $n(\mathbf{x})$, and dispersion $\mathbf{D}'_\psi(\mathbf{x})$. The FPME is

$$n(\mathbf{x})[u\tilde{c}(\mathbf{x}, u) - c_0(\mathbf{x})] = -\nabla \cdot \Big\{\tilde{M}(u; \mathbf{x})\mathbf{q}_\psi(\mathbf{x})\,\tilde{c}(\mathbf{x}, u) - \nabla \cdot \Big[\tilde{M}(u; \mathbf{x})\mathbf{D}'_\psi(\mathbf{x})\,\tilde{c}(\mathbf{x}, u)\Big]\Big\}. \quad (71)$$

(Note that in (71), $\mathbf{D}'_\psi(\mathbf{x})$ is a factor for the concentration $\tilde{c}$, whereas in the classical ADE approach, $\mathbf{D}'_\psi(\mathbf{x})$ is a factor for $\nabla\tilde{c}$.) Developing the second derivative in (71) we can write

$$n(\mathbf{x})[u\tilde{c}(\mathbf{x}, u) - c_0(\mathbf{x})] = -\tilde{M}(u; \mathbf{x}) \cdot \Big\{\Big[\mathbf{q}_\psi(\mathbf{x}) - \nabla \cdot \mathbf{D}'_\psi(\mathbf{x})\Big] \cdot \nabla\tilde{c}(\mathbf{x}, u) - \nabla \cdot \Big[\mathbf{D}'_\psi(\mathbf{x}) \cdot \nabla\tilde{c}(\mathbf{x}, u)\Big] + \tilde{c}(\mathbf{x}, u)\nabla^2\mathbf{D}'_\psi(\mathbf{x})\Big\}. \quad (72)$$

[140] For the hydrogeological applications considered herein, it can be shown [*Cortis et al.*, 2004b] that spatial derivatives of $\tilde{M}(u; \mathbf{x})$ are small and are not included in (72); we shall assume that the term $\nabla^2\mathbf{D}'_\psi(\mathbf{x})$ can be neglected, so that (72) simplifies to

$$n(\mathbf{x})[u\tilde{c}(\mathbf{x}, u) - c_0(\mathbf{x})] = -\tilde{M}(u; \mathbf{x}) \cdot \Big\{\Big[\mathbf{q}_\psi(\mathbf{x}) - \nabla \cdot \mathbf{D}'_\psi(\mathbf{x})\Big] \cdot \nabla\tilde{c}(\mathbf{x}, u) - \nabla \cdot \Big[\mathbf{D}'_\psi(\mathbf{x}) \cdot \nabla\tilde{c}(\mathbf{x}, u)\Big]\Big\}. \quad (73)$$

Dropping the term $\nabla^2\mathbf{D}'_\psi$ is reasonable when the $\mathbf{D}'_\psi$ field can be seen as varying piecewise. Of course, the term should not be neglected if the dispersion field can be resolved on intermediate scales smaller than the large (facies) scale (but larger than the microscopic heterogeneities). Clearly, for a macroscopically homogeneous medium, i.e., with no variations in n, $\mathbf{q}_\psi$, $\mathbf{D}'_\psi$, and $M(u)$ over the scale $\mathbf{x}$, then (73) reduces to the form (33) but at the scale $\mathbf{x}$. Recall that $M(u)$ as given by (29) accounts for the small-scale ($\mathbf{y}$) heterogeneities.

[141] The effects of small-scale heterogeneities are reflected in the memory term $\tilde{M}(u; \mathbf{x})$ which gives rise to anomalous (non-Fickian) dispersion. The deterministic large-scale variations of the heterogeneous medium are reflected by $\mathbf{q}_\psi(\mathbf{x})$ and $\mathbf{D}'_\psi(\mathbf{x})$. In the Fokker-Planck equation the effect of the macroscopic heterogeneities are additionally included in the important "drift correction" term, $\nabla \cdot \mathbf{D}'_\psi(\mathbf{x})$, which modifies $\mathbf{q}_\psi(\mathbf{x})$ on the right side of (73).

4.3. Numerical Simulation

[142] We now illustrate how the above framework can be applied to treatment of a nonstationary, heterogeneous system, e.g., actual field sites. Here we demonstrate the versatility and computational feasibility of the approach.

[143] We consider transport in the two-dimensional domain Ω, depicted in Figure 31. The unit (1×1) domain is subdivided into macroscopically "homogeneous" regions of three different materials Ω_1, Ω_2, and Ω_3 whose porosity, permeability, and dispersivity are given in Table 1. For the purposes of this set of simulations the $\alpha_\psi(\mathbf{x})$ ((31) but with $\mathbf{D}'_\psi(\mathbf{x})$ instead of $\mathbf{D}_\psi$) is assumed to be known from field measurements, as are all the spatially dependent parameters in $\tilde{M}(u, \mathbf{x})$.

[144] We impose constant pressure BCs equal to 1 and 0 on the inlet and outlet boundaries, respectively, and a no-flow condition over the remaining portions of the boundary. We assume a zero initial resident concentration c_0, a constant concentration at the injection point (0.1,0.1), a free BC at the outlet boundary, and a no-flux BC on the remainder of the domain. The approximation of the interface zone between the different macroscopically homogeneous regions is obtained by linearizing the spatial derivative of the dispersivity over a small region about the interface itself.

[145] The first step in the actual computations is the solution of the flow field $\mathbf{q}(\mathbf{x})$ by means of any classical solver. For simplicity, we make the approximation $\mathbf{q}(\mathbf{x}) \sim \mathbf{q}_\psi(\mathbf{x})$, discussing it further at the end of this section. The second step is the solution of (73) for all values of the u variable, using the algorithm of *de Hoog et al.* [1982]. The final step is a numerical $\mathcal{L}^{-1}$ of the Laplace transformed solutions for every point of interest in space.

[146] As discussed at the beginning of section 3, solution of (73) requires specification of the transition rate probability density function $\psi(t)$. Clearly, for application to specific, real systems the functional form of $\psi(t)$ and parameter values within it must be fit or derived from measurable properties of the medium, the flow field, and/or the tracer transport itself. Here we shall consider the truncated power law form (61); other forms of $\psi(t)$ are considered by *Cortis et al.* [2004b].

[147] Within the context of this $\psi(t)$ function we employ two sets ($\mathcal{B}$ and $\mathcal{S}$) of β values in the regions Ω_1, Ω_2, and Ω_3 (see Table 2). In set $\mathcal{B}$ the β coefficients are larger than 1, whereas in set $\mathcal{S}$ they are smaller than 1. We recall that $\beta < 1$ is typical of highly dispersive (non-Fickian) transport, while $1 < \beta < 2$ is associated with moderately dispersive systems;

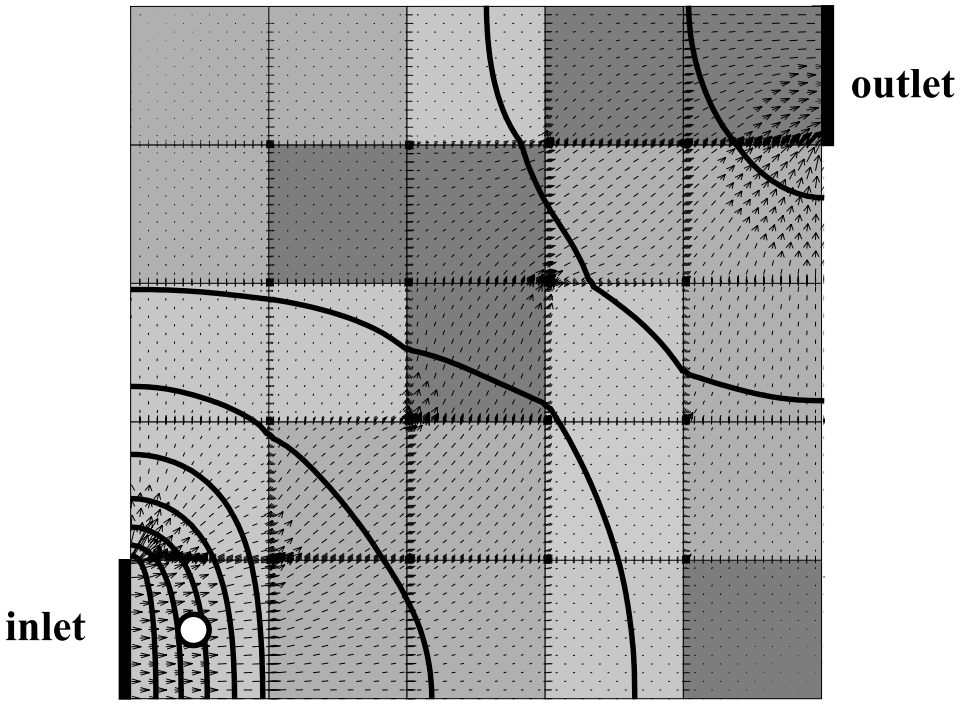

Figure 31. Nonstationary domain used for the computations. The three regions Ω_1, Ω_2, and Ω_3 are represented in the light, medium, and dark gray, respectively. The overall region spans 1×1 dimensionless units. The point source is located at the coordinates (0.1, 0.1). The inlet and outlet regions are also indicated. After *Cortis et al.* [2004b].

$\beta > 2$ leads to Fickian transport. For each set $\mathcal{B}$ and $\mathcal{S}$ of β coefficients, tracer transport was calculated for the FPME. For comparison, tracer behavior was also determined from application of the standard ADE model, i.e., by using (73) and setting the memory term to unity and the drift correction term $\nabla \cdot \mathbf{D}_\psi(\mathbf{x})$ to zero. Recall that the definitions of dispersivity in Table 1 are different with respect to the FPME and the ADE. More specifically, α_ψ as it is used in calculations with the FPME is given by (32) (but with $p(\mathbf{s}; \mathbf{x})$ instead of $p(\mathbf{s})$). In contrast, α as applied to the classical ADE is the conventional "dispersivity" coefficient.

[148] Figure 32 shows contour plots of the migration of tracer plumes through the domain shown in Figure 31, as given by solution of the FPME for the truncated power law $\psi(t)$ function (61), with $\beta > 1$ (set $\mathcal{B}$, see Table 2), along with that given by solution of the ADE. Three time frames, for "early," "intermediate," and "late" times, are reported for the sake of comparison. The choice of $\beta > 1$ was made so as to yield timescales for the rate of tracer advance that are of the same order of magnitude as those obtained from solution of the ADE. As discussed above, this range of β values is applicable to systems with moderate dispersivity properties. Deviation of tracer transport behavior from that given by the ADE is apparent at all times. More specifically, a retardation of the tracer at all times is evident, because the exponential cutoff time t_2 is much larger than t_{max}, the maximum time spread from first tracer arrival to last arrival for the considered parameters. Note also that when compared to the ADE solution in Figure 32b, the isoconcentration contours in Figure 32a are "compressed" along the interfaces of maximum dispersivity contrast. This is due to the effect of the "drift correction" term mentioned at the end of section 4.2.

[149] We consider in Figure 33 the FPME solutions for the truncated power law $\psi(t)$ function, with $\beta < 1$ (set $\mathcal{S}$). Note that relative to Figure 32 the timescales for the overall rate of tracer advance increase by an order of magnitude

TABLE 1. Values of the Permeability k, Longitudinal and Transverse Dispersivity α_ψ^L and α_ψ^T and Porosity n for the Three Homogeneous Facies Ω_1, Ω_2, and Ω_3

	Ω_1	Ω_2	Ω_3
k_x	0.10	0.50	0.40
k_y	0.10	0.50	0.40
$\alpha_\psi{}^L$	0.05	0.025	0.10
$\alpha_\psi{}^T$	0.005	0.0025	0.01
n	0.30	0.35	0.40

TABLE 2. Values of Coefficients for the Transition Probability Function $\psi(t)$ in Equation (61) for the Three Homogeneous Facies Ω_1, Ω_2, and Ω_3 for Cases $\mathcal{B}$ and $\mathcal{S}$

	Ω_1	Ω_2	Ω_3
	Case $\mathcal{B}$ ($\beta > 1$)		
t_1	0.10	0.15	0.20
t_2	1×10^6	5×10^6	10×10^6
β	1.25	1.35	1.45
	Case $\mathcal{S}$ ($\beta < 1$)		
t_1	0.10	0.15	0.20
t_2	1×10^6	5×10^6	10×10^6
β	0.60	0.70	0.80

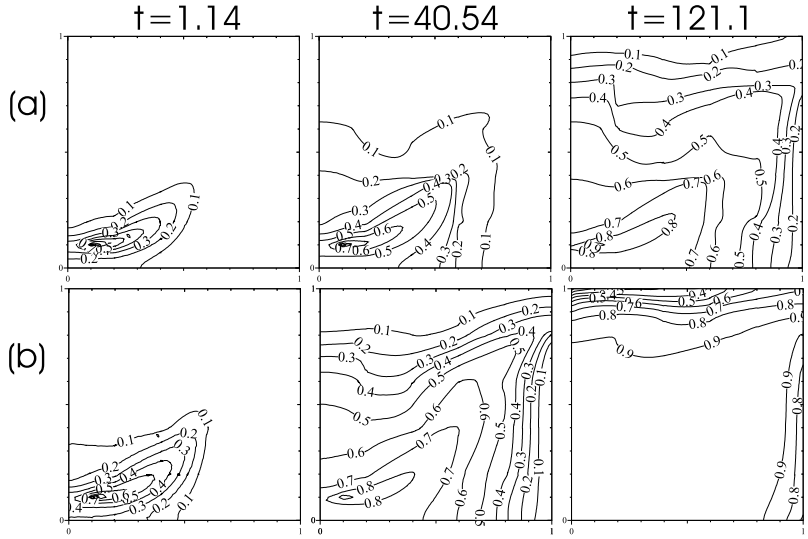

Figure 32. Contour plots of tracer concentration for plume migration through the heterogeneous domain shown in Figure 31. Shown here are solutions of (a) FPME with the truncated power law $\psi(t)$ given by (61), with set $\mathcal{B}$ parameters (Table 2), and (b) ADE. After *Cortis et al.* [2004b].

(and thus the contours for the ADE solution are totally different; see Figure 32b). In fact, although the actual contour patterns are totally different, the early time plot shown in Figure 33 corresponds to the late time plot shown in Figure 32. This dramatic effect is due to the highly anomalous dispersion characteristics of media with values of $\beta < 1$. This a key point of our numerical results using the hybrid method. Even if just one region in a field (with large heterogeneity) has $\beta < 1$, the differences with the conventional numerical results are significant (in the timescale and the contour configurations); that is, the residues impact the large-scale transport [*Eggleston and Rojstaczer*, 1998].

[150] Considering Figure 32a and the corresponding BTC obtained by integrating the tracer concentration flux over the outlet boundary (not shown here) [see *Cortis et al.*, 2004b], the time for the relative concentration to reach unity is of the order of 10^2, while the value of t_2 is of the order of 10^6. Here the tracer reacts strongly to heterogeneities in the domain, relative to the overall transport length scale. In other words, when the truncation time is much larger than the overall transport time, transport remains highly non-Fickian, and the domain is not "homogenized." A detailed analysis of the effects of t_2, relative to the overall transport time, is given by *Dentz et al.* [2004] (see section 3.7). In contrast, for the simulation based on $\beta < 1$ (Figure 33), the cutoff time t_2 is roughly of the same order of magnitude as the overall breakthrough time. In such cases, transport evolves from non-Fickian to Fickian, as the domain becomes "homogenized" relative to the migrating tracer.

[151] For application of this approach to actual field situations a few comments are in order. Estimates of the parameters in the $\psi(t)$ function, as well as of α_ψ, can be obtained by fitting the CTRW solution to representative BTCs measured in each of the different facies, similar to estimates of the familiar (but incorrect!) dispersivity of the ADE. In addition, we recall that, in general, $\mathbf{v}_\psi$ as it appears in (33) is not equal to the fluid velocity $\mathbf{v}$. At least, in principle, $\mathbf{v}_\psi = \bar{\mathbf{s}}/\bar{t}$ can be estimated by determining the characteristic time of travel $\bar{t}$ across $\bar{\mathbf{s}}$, where $\bar{\mathbf{s}}$ is the mean of $\mathbf{s}$ [e.g., *Bijeljic and Blunt*, 2006]. In practice, this tracer velocity can be estimated directly from measured tracer BTCs, simultaneously with the estimates of the aforemen-

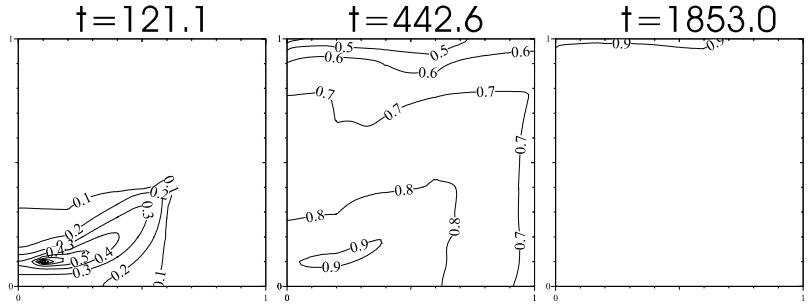

Figure 33. Contour plots of tracer concentration for plume migration through the heterogeneous domain shown in Figure 31. Shown here are solutions of FPME with the truncated power law $\psi(t)$ given by (61), with set $\mathcal{S}$ parameters (Table 2). After *Cortis et al.* [2004b].

tioned parameters. Of course, one generally calculates only the Darcy velocity field on the basis of an estimated hydraulic conductivity field. One possibility is therefore to modify the estimated velocities in proportion to estimated ratios between $\mathbf{v}$ and $\mathbf{v}_\psi$. On the other hand, the uncertainty in estimates of the fluid velocity field $\mathbf{v}$ due to uncertainties in estimates of the hydraulic conductivity field make the approximation $\mathbf{v}_\psi \approx \mathbf{v}$ both convenient and reasonable. The full expression of the tracer velocity is given by (28), i.e., including the factor of the memory function $\tilde{M}(u)$, and hence is different from $\mathbf{v}$. As a side consideration we note that tracer tests intended to estimate the hydraulic conductivity between two wells, in fact, provide an estimate only of the tracer velocity rather than the actual hydraulic conductivity between the wells.

5. CRITIQUE OF THE CTRW APPROACH

[152] The applicability of the CTRW approach to a range of experiments was shown in section 3, and the numerical simulations of transport in nonstationary domains (section 4) demonstrate further applicability of CTRW to field situations.

[153] These features indicate that the interactions between the transporting particle and the medium can be mapped effectively onto an appropriate choice of $\psi(\mathbf{s}, t)$. Furthermore, a few characteristics of the $\psi(\mathbf{s}, t)$ are very often sufficient to capture completely the particle dynamics. This description begs the immediate question of how one obtains $\psi(\mathbf{s}, t)$. A considerable effort may be required, but one can proceed in a hierarchy of levels of approximation. Here we suggest five possible approaches: (1) Fit the measurements with a simple form of $\psi(\mathbf{s}, t)$, e.g., (45), and use β as a fitting parameter. (2) Develop a ''library'' of $\psi(\mathbf{s}, t)$ for different types of geological formations and flow conditions. (3) Obtain the velocity histogram for a RFN [*Scher et al.*, 2002a], a porous medium [*Bijeljic and Blunt*, 2006], or permeability field [*Di Donato et al.*, 2003] and model $\psi(\mathbf{s}, t)$ in the manner of (40). (4) Perform a numerical simulation on part of a complex system, e.g., an array of intersecting fracture platelets, and determine $\psi(\mathbf{s}, t)$ in terms of variables such as platelet size and aperture and then use a pdf of these variables to develop $\psi(\mathbf{s}, t)$ for the entire system. (5) Determine the $w(\mathbf{s}, \mathbf{s}')$ and calculate the ensemble average as in (18). Also, in the case of a multiple trapping scenario, which is discussed in detail in section 7.1, constraints on $\psi(\mathbf{s}, t)$ can be obtained by using its specific connection to mass transfer.

[154] Explicit and analytical linking or conditioning of $\psi(\mathbf{s}, t)$ to known physical information, such as the heterogeneity of the hydraulic conductivity field, is certainly a key area to develop. A step in this direction is given by *Dentz and Berkowitz* [2005], who demonstrate a $\psi(\mathbf{s}, t)$ that is based explicitly on the underlying heterogeneity distribution, for the case of transport under spatially random adsorption. On the other hand, it should be recognized that there naturally are limitations to full determination of $\psi(\mathbf{s}, t)$, or for that matter any such effective transport description, from purely theoretical considerations. However, to reiterate, on a practical level such full determination of $\psi(\mathbf{s}, t)$ is not crucial. The remarkable feature documented in this paper is that a very few parameters capturing the important features of $\psi(\mathbf{s}, t)$ are sufficient to account quantitatively for a host of observations. Even the well-known determination of the macrodispersion coefficient given by *Gelhar and Axness* [1983], as used in the ADE (see section 6.2.1), is predicated on the assumption of Fickian transport but also on ignoring completely the dependence of this parameter on the cutoff (transition) time between non-Fickian and Fickian transport (recall section 3.7.3). Moreover, a priori prediction of this transition is not trivial, being dependent on, e.g., the velocity distribution, the particle residence time, the heterogeneity scale of the hydraulic conductivity, and the boundary conditions. Thus full specification of a transport equation and parameter values must ultimately rest on site-specific measurements. As noted in the preceding paragraph, several approaches can be taken to obtain this information. In particular, because the underlying permeability distribution of a domain gives rise to the velocity distribution, which thus includes naturally all correlations affecting transport, the velocity distribution can be used to define the actual particle distributions.

[155] Treatment of many sorption and other multiple trapping pictures is straightforward, as discussed in section 7.1 and shown by, e.g., *Margolin et al.* [2003]. Further extension of the CTRW formulation, i.e., specification of either a $\psi(\mathbf{s}, t)$ and/or modifications to, e.g., the transport equation (33), to account explicitly for biogeochemical reactions such as precipitation/dissolution or other, often nonlinear (feedback), reactions is another area for future research. One important step of this CTRW extension is taken at the end of section 7.1, where an interaction between a dispersive medium and a spectrum of mass transfers into immobile states is developed.

[156] Another important area for further development focuses on methods of solving the transport equation. While the mathematics are somewhat more extended than those involved with the familiar ADE, CTRW solutions are accessible, and the pde formulation (section 2.6) enables adaptation of many existing solution techniques. This review has focused on the methodology of Laplace transforms, but one can work directly in the time domain, if desired, using (35).

[157] Because CTRW results from an ensemble average, the concentration distribution in CTRW is also an ensemble-averaged quantity, which we usually assume to be sufficient for a system size large compared to the scale of heterogeneity. The variation in the concentration, at a specific location, among different realizations of the underlying random fields is thus not quantified explicitly; that is, the average concentration is an ensemble average. In other words, if the system size L is much greater than the scale of heterogeneity l, then the rearrangement of the heterogeneity disorder from one realization to the next will result in a small change in a given neighborhood. Hence the variation from the ensemble-averaged $C(\mathbf{s}, t)$ is expected to be small. Furthermore, the dynamics of $c(\mathbf{s}, t)$ are governed

by $\psi(\mathbf{s}, t)$, which is a pdf based on the flow field (and thus heterogeneity distribution) of the entire system; this feature enhances the sampling of each locale to the heterogeneity distribution of the whole system. In contrast, if $L \ll l$, then the variations in realization-to-realization configurations can lead to large heterogeneity differences in a given locale and therefore large variations in concentration. For systems with $L \gg l$ it is best to use the hybrid approach; the ensemble average is then used for the small-scale residual heterogeneities. Thus the issue of variations is closely tied to the comparison between L and l. On a physical basis the ensemble average is best utilized when $L > l$.

[158] On a formal basis, "predictive uncertainty" in the CTRW can be considered on two levels. The first level involves the GME (2), which is an equation for the ensemble-averaged $C(\mathbf{s}, t)$. In principle, an equation for the variance of this function can also be developed from the master equation itself. The second level is a practical one: How sensitive are the results to variations in the parameters that one inputs into $\psi(\mathbf{s}, t)$? In any application with a model $\psi(\mathbf{s}, t)$ one can easily use Monte Carlo simulations to compute the effects of these variations on final predictions and hence assess the "predictive uncertainty." It remains to examine these issues in detail.

6. ADVECTION-DISPERSION EQUATION AND UPSCALING

[159] In section 1 we discussed the intensive efforts to model transport in porous media over the last decades. Many of these efforts are basically tied to the ADE. In sections 6.1 and 6.2 we discuss in detail the ramifications of the ADE approach and related modeling frameworks. We then contrast them to the CTRW framework with a particular emphasis on the different methods of averaging of the disorder-induced fluctuations of transport quantities.

[160] The ADE is used extensively in all the natural sciences; for example, in semiconductor physics it describes the flow of electrons due to an applied voltage and allows for diffusive motion driven by spatial variations of the electron density. In liquid transport the ADE is used, e.g., to calculate Taylor dispersion (molecular diffusion of particles in a flowing fluid in a pipe). The ADE is

$$\frac{\partial c(\mathbf{s}, t)}{\partial t} = -\mathbf{v} \cdot \nabla c(\mathbf{s}, t) + \mathbf{D} : \nabla\nabla c(\mathbf{s}, t), \tag{74}$$

which is the same as (27) for constant fluid velocity $\mathbf{v}$ and dispersion $\mathbf{D}$. Hence one derivation of the ADE applied to porous media can be based on the kinetics described by the master equation (1) with a Taylor expansion of $C(\mathbf{s}, t)$, (23), and $w(\mathbf{s}, \mathbf{s}')$ [*Berkowitz et al.*, 2002].

[161] The classical derivation of the ADE for porous media [e.g., *Bear*, 1972] is based on the assumption of the existence of a representative elementary volume (REV), i.e., on the assumption that at some scale $x \gg y$ the variations of the porous medium at the x scale can be considered homogeneous. The other required assumptions are that (1) the porous medium is fully saturated; (2) Darcy's law applies; (3) transport of a tracer can be split a priori into an advective part and a mechanical dispersion part; (4) mechanical dispersion obeys Fick's law, where the coefficient $\mathbf{D}$ is assumed to be composed of a molecular diffusion part, D_m, and a velocity-dependent part, which in one-dimensional form is written $D = D_m + v\alpha$, with α the so-called dispersivity usually assumed to be a characteristic length of the pores inside the REV; (5) the transport velocity equals the fluid velocity; and (6) the spatial variation of the fluid velocity inside the REV can be neglected.

[162] The historical motivation for the REV stems from the need to use a continuous mechanistic approach for porous media, i.e., media which are inherently discontinuous at the scale of the pores. For some quantities this is a very useful notion, as in defining geometrical properties like porosity, specific surface, and permeability as an average over the REV. These quantities can in some sense be considered "local quantities." Averaging over the REV to define other quantities such as dispersivity is more limited and is a point of departure from the CTRW approach, as we discuss in section 6.1.

[163] It must be recognized that use of the ADE, at a variety of different scales, is a key aspect of the vast majority of transport theories. The dispersivity α then takes on the role of a "scaling" parameter in the sense that varying its magnitude is the basis for fitting ADE solutions to BTC measurements ranging over several orders of magnitude. A major drawback of the approach is that, in contrast to the fundamental assumption that α is an intrinsic porous medium constant, "force fitting" the ADE to measurements over scales from the small laboratory column to the large field scale demonstrates an ad hoc "space dependence" or "time dependence" of α [e.g., *Lallemand-Barres and Peaudecerf*, 1978; *Gelhar et al.*, 1992].

[164] A number of approaches have been made to connect the different scales and not just use (74) with $\mathbf{D}$ or α as a fitting parameter. All of these approaches start from the assumption that the ADE equation holds at some microscopic scale and then assume further that some kind of macrodispersion parameter can be found starting from the information contained at a microscopic scale. In many approaches it is also assumed that the ADE form holds at each scale with upscaled coefficients. The two main approaches, which to some degree overlap, we denote for brevity as "averaging" and "stochastic" methods; we discuss them in detail in sections 6.1 and 6.2.

6.1. Volume Averaging

6.1.1. Methods

[165] The aim of averaging is to start with the ADE at the microscopic scale and move to a larger scale to obtain an ADE with modified coefficients. The challenge is to account for the deviation from the mean, due to the microscopic fluctuations, of the velocity and dispersion. We will highlight a few of these different approaches.

[166] In "volume averaging," moving up from one scale to another in a porous medium necessitates the use of some

kind of average operator on the microscopic fields at the scale L_y to obtain fields which are significant at the macroscopic scale L_x. This average operator acts on some volume of the porous medium and assigns a macroscopic value to it. Therefore, unlike the microscopic fields, the macroscopic fields are smeared over space.

[167] The question then arises: What is the "correct" size of the averaging volume? The usual answer is to define a REV as the smallest volume of integration for which there are no fluctuations in the averages of one or more of the characteristic features of the porous medium; for example, *Bear* [1972] defines the REV in terms of porosity. The size of the REV, L_x, must be intermediate between a characteristic microscopic length and a characteristic macroscopic length scale of the sample, L_R, i.e., $L_y \ll L_x \ll L_R$. Sometimes these length scales depend not only on the microstructure of the porous medium but also on the physical process under study. It is also possible that porous media with clear separations of the geometric scales do not have a REV for a particular physical process. We argue that this is the case for the dispersion problem in porous media. Moreover, *Berkowitz and Bachmat* [1987] showed that a REV-like dispersion tensor is necessarily scale (REV size) dependent when deviations of the macroscopic velocity are taken into account.

[168] A general definition of the volume average operator, indicated by angle brackets, can be made in terms of convolution products of spatial distribution functions [*Cushman*, 1984; *Quintard and Whitaker*, 1994]. An important result relating the gradient of some quantity φ at the microscales and macroscales is represented by a theorem that states that the average of the gradient of φ, $\langle \nabla \varphi \rangle$, at the microscopic scale y equals the gradient at the macroscopic scale x of the averaged physical quantity $\nabla \langle \varphi \rangle$ plus a fluid-solid surface average contribution.

[169] A modified accounting for the role of deviations from average values, which arise from the same upscaling framework, is to use the "volume averaging with closure" method. Here the term "closure" refers to a particular integrodifferential problem, which determines definition of the deviations of the quantities of interest from their mean value using a pde. This method thus attempts to assign coefficients that prescribe the functional dependence of deviations from average quantities on the basis of a physically meaningful microscopic process. A detailed account of the approach is given by *Whitaker* [1999]. The basic requirements of this method are (1) separation of the microscopic and macroscopic length scales (see discussion below); (2) periodicity on the boundaries of the REV; (3) a phenomenological (postulated) relationship between the deviations of the quantities of interest from the average value and the gradient of the average quantity itself; (4) a coefficient of proportionality in requirement 3 that usually satisfies an integrodifferential problem that is similar to, or can be mapped onto, the original microscopic equations; and (5) a starting microscale equation that is invariably an ADE.

[170] A typical definition of the macrodispersion coefficient (for constant porosity n and dropping the bold notation for vectors) is given by

$$\frac{\partial \langle c \rangle}{\partial t} + \langle v \rangle \frac{\partial \langle c \rangle}{\partial s} - \frac{\partial}{\partial s}\left(\mathcal{D} \frac{\partial \langle c \rangle}{\partial s} \right) = 0 \tag{75a}$$

$$\mathcal{D} = D_m + D_m \frac{1}{V_f} \int_{A_{fs}} (b_f - b_s) n_{fs} dA - \langle (v - \langle v \rangle) b \rangle, \tag{75b}$$

$$(c - \langle c \rangle) = b \frac{\partial \langle c \rangle}{\partial s}, \tag{75c}$$

where the vector b is given by the problem

$$v \frac{\partial b}{\partial s} - \frac{\partial}{\partial s}\left(D_m \frac{\partial b}{\partial s} \right) = -(v - \langle v \rangle) \tag{76a}$$

$$\frac{\partial b}{\partial s} = -I \quad \text{on} \quad A_{fs} \tag{76b}$$

$$b = 0, \quad \text{for} \quad t = 0 \tag{76c}$$

$$b(s + \vartheta) = b(s), \tag{76d}$$

where (76d) is periodicity, I is the identity tensor, A_{fs} denotes the fluid-solid surface, V_f is the fluid volume, n_{fs} is the normal vector to the fluid-solid surface, and the subscripts f and s denote fluid and solid phases, respectively. Thus the uncertainty in the fluctuations from the mean is mapped onto the quantity b, which, in turn, solves a microscopic ADE problem. The requirement of periodicity at the scale of the REV is key to this method. *Bhattacharya and Gupta* [1990] proved that to derive the ADE (in a central limit theorem framework, for instance), the requirement of periodicity (or quasiperiodicity) on the fluid velocity v is needed.

[171] "Homogenization theory" [see, e.g., *Rubinstein and Mauri*, 1983; *Bourgeat et al.*, 1988] and the "renormalization group" are upscaling techniques that have been used in the context of volume averaging. These approaches require an underlying REV at some scale for which a significant average concentration can be defined locally. Locality is enforced by the specification of periodic BCs on the relevant quantities over the local cell (REV). In these theories an important parameter for the upscaling of the equations is $\epsilon = L_y/L_x$, where L_x is the length scale of the macroscopic variations and L_y is the length scale corresponding to the local variations.

[172] The homogenization method establishes a hierarchy of equations based on an expansion of the original transport problem in ε. A closure hypothesis is needed to truncate the hierarchy; the usual one chosen is a Fickian one and is analogous to (75c) [see, e.g., *Bourgeat et al.*, 1988]. Rigorous derivation of a stochastic homogenization method that does not require periodic BCs is given by *Auriault and Adler* [1995] and *Lunati et al.* [2002].

[173] "Renormalization group" techniques have been used in the context of subsurface hydrology to determine macrodispersion coefficients for solute transport in random flow fields [e.g., *Koch and Shaqfeh*, 1992; *Zhang*, 1995; *Jaekel and Vereecken*, 1997]. Starting from an ADE with random advection, the macrodispersion coefficients are expanded into a perturbation series in the fluctuations of the random flow field. The renormalization group represents in this context a tool to systematically sum up certain contributions of the perturbation series (and in its simplest application a so-called one-loop renormalization is used). The resulting macrodispersion coefficients are beyond second-order perturbation theory in the random field fluctuations, similar to the results obtained from the application of "Corrsin's conjecture" [e.g., *Dagan*, 1994b; *Zhang*, 1995]. The latter, however, has been shown to be inconsistent in two dimensions [*Dagan*, 1994b; *Dentz et al.*, 2003; *Attinger et al.*, 2004].

6.1.2. Length-Scale and Timescale Separation: A Critical Discussion of the ADE and Averaging

[174] At the start of section 6 we noted basic assumptions required for applicability of the ADE. The assumptions that a porous medium can be considered homogeneous at the relevant scale of measurement and that transport mechanisms can be separated a priori into components of advection and hydrodynamic dispersion are highly restrictive. However, as discussed in section 1, illustrated in the "homogeneous" medium shown in Figure 1, and shown consistently throughout the examples in section 3, the key underlying assumption that small fluctuations can be neglected is clearly inadequate. Thus ADE-type descriptions of tracer transport and use of, e.g., REV approaches are rarely fully correct even on local scales of several centimeters.

[175] At best, ADE-type descriptions usually capture only the average properties of tracer migration. Only when the transport length is orders of magnitude larger than the heterogeneity scale does homogenization occur, with the result that Fickian transport is indeed present in the system. For real domains with a finite hierarchy of heterogeneity scales, but also finite (and usually limited) domain lengths, i.e., for the vast majority of field-scale transport problems of interest, a REV-like homogenization limit does not exist. To make matters more complex, determination of the "cutoff" at which Fickian transport descriptions are correct is far from rigorous. As shown in the examples in section 3.7.3, Figures 24–26, tracer plumes can retain a non-Fickian behavior long after the transport time/length scale is Fickian because of memory effects on the evolution of the tracer plume up to the cutoff.

[176] Let us consider these issues further, recognizing that an efficient and meaningful upscaling to large spatial and temporal scales is required of any practical transport theory. As shown by *Levy and Berkowitz* [2003] (e.g., Figure 9), analysis of solute transport in "macroscopically homogeneous" media indicates that flow and transport do not "homogenize" on the same temporal and spatial scales. The fundamental importance of time and length scales relative to the scales of heterogeneity and domain size were also examined in the context of the two column experiments in section 3.5. The transition from non-Fickian to Fickian behavior as the column length increased showed that the difference in residence time was the determining factor. In other words, the validity of Darcy's law on certain spatial scales does not automatically imply that transport of a passive solute can be described by the standard Fickian theory.

[177] A consequence of this fact is that the effects of transport processes cannot simply be separated into "independent" mechanisms. Recall, for example, the discussion in section 3.4.2 and Figure 11. Because the timescale changes as the overall flow rate changes, the transport parameters (e.g., β or equivalent parameters for other forms of $\psi(t)$ in the CTRW framework) are not "intrinsic" and constant (see Figure 3). This result is in stark contrast to the specification of precisely these assumptions for, e.g., the dispersivity α in the ADE. This behavior accounts also for the intertwining of the two dispersion mechanisms in the CTRW-based FPME formulation (71) in contrast to the usual ADE or stochastic approaches (see section 6.2), which attribute the spreading of the BTC only to the second spatial moment of the tracer distribution. The memory function (29) is indicative of a nonlocal-in-time dispersion, whereas the $\mathbf{D}_\psi$ parameter (31), required to fit the entire BTC, provides a measure of the local-in-space dispersion.

[178] Traditional transport theories have focused on spatial heterogeneity. A key feature of the CTRW approach, on the other hand, is the emphasis on temporal aspects of particle transport, induced by spatial heterogeneity. Shifting the focus to work within the CTRW framework therefore represents a change in paradigm.

[179] As discussed in section 6.1.1, some of the averaging techniques require that the porous medium be periodic (i.e., the hypothesis of periodicity is a crucial one in the argument of *Brenner* [1980], together with the hypothesis of ergodicity). While assuming periodicity is a convenient idealization for small-scale fluid flow applications [see, e.g., *Dorfman and Brenner*, 2002, and references therein], specification of natural geological structures as periodic is not an appropriate starting point for modeling contaminant migration. Another consequence of the periodicity requirement should also be recognized: *Brenner* [1980] showed that for periodic porous media the "fluid velocity" $\mathbf{v}$ is necessarily identical to the "transport velocity" $\mathbf{v}_\psi$ and traced the difference between the two velocities only to some particle-size dependent exclusion effect. While this is certainly an important factor, for instance, in colloidal transport (where dispersion generally plays a minor role), other mechanisms are present as well.

[180] CTRW theory, on the other hand, is not limited by the assumption of periodicity (nor of ergodicity). Periodicity is, in fact, strictly forbidden at a local scale, because the structure of the master equation (1) requires, in general, an accounting of all jumps in the domain. Relaxing these assumptions in a CTRW framework leads to the correct physical picture that the fluid velocity $\mathbf{v}$ is, in general,

different from the tracer transport velocity $\mathbf{v}_\psi$. In the CTRW picture the relative weights of each jump, and therefore the possibility for a tracer to explore different positions in space, are governed by the transition times w^{-1}.

6.2. Stochastic Approach

[181] Rather than give a comprehensive review, in this section we highlight two important aspects of the stochastic approach to subsurface hydrology that are of relevance in the present context: (1) non-Fickian transport and (2) effective transport descriptions and their relation to CTRW. We emphasize that the term "stochastic approach" is a general one and can, in principle, include a wide variety of formulations (including CTRW) and solution methods. However, we shall use the term as it has been identified in recent years, namely, with the particular framework outlined below.

[182] The stochastic approach has been considered extensively in subsurface hydrology and applied to a wide variety of transport situations ranging from passive and reactive solute transport to the study of seawater intrusion in spatially heterogeneous environments. For thorough overviews of stochastic modeling in hydrology we refer the interested reader to the textbooks by, e.g., *Dagan* [1989], *Gelhar* [1993], *Dagan and Neuman* [1997], *Zhang* [2002], and *Rubin* [2003]. Critical reckoning of the current position of stochastic modeling in subsurface hydrology can be found in a series of papers [*Christakos*, 2004; *Dagan*, 2004; *Freeze*, 2004; *Ginn*, 2004; *Molz*, 2004; *Neuman*, 2004; *Rubin*, 2004; *Sudicky*, 2004; *Winter*, 2004; *Zhang and Zhang*, 2004].

[183] In a stochastic approach, spatially and temporally fluctuating system parameters such as hydraulic conductivity, porosity, and chemical properties of the medium, for example, are modeled as random fields characterized by specific, experimentally accessible statistical properties. The effective transport behavior of a solute can then be obtained by ensemble averaging of the observables of interest over all realizations of the respective random fields.

6.2.1. Transport Coefficients and Non-Fickian Behavior

[184] The influence of medium heterogeneities on large-scale transport can be quantified in terms of effective transport coefficients such as the effective center of mass velocity and effective dispersion coefficients, which are derived from the first and second moments of the normalized concentration $c(\mathbf{s}, t)$, respectively. The transport equations for $c(\mathbf{s}, t)$ are discussed in section 6.2.2.

[185] In the stochastic approach, large-scale transport coefficients are defined as averages over all possible realizations of the respective random fields. We focus here on "ensemble dispersion" coefficients $D_{ij}^{\mathrm{ens}}(t)$ [e.g., *Kraichnan*, 1959; *Roberts*, 1961; *Gelhar and Axness*, 1983; *Neuman et al.*, 1987; *Dagan*, 1984, 1988], which are derived from the ensemble-averaged concentration distribution, $c(\mathbf{s}, t)$:

$$D_{ij}^{\mathrm{ens}}(t) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left[\int s_i\, s_j\, c(\mathbf{s},t)\,\mathrm{d}^d s - \int s_i\, c(\mathbf{s},t)\,\mathrm{d}^d s \int s_j\, c(\mathbf{s},t)\,\mathrm{d}^d s\right]. \tag{77}$$

Note that the $D_{ij}^{\mathrm{ens}}(t)$ characterize the spreading of the average solute distribution and not necessarily solute spreading in a typical heterogeneity realization, as opposed to "effective" dispersion coefficients [e.g., *Batchelor*, 1949; *Kitanidis*, 1988; *Dagan*, 1990, 1991; *Rajaram and Gelhar*, 1993; *Zhang and Zhang*, 1996; *Attinger et al.*, 1999; *Dentz et al.*, 2000]. The $D_{ij}^{\mathrm{ens}}(t)$ is independent of time if, e.g., the concentration distribution is Gaussian. We focus on $D_{ij}^{\mathrm{ens}}(t)$ because it is related to the dispersive flux in an effective upscaled transport equation for the average solute concentration, as outlined in section 6.2.2.

[186] A principal result of the stochastic approach is that of *Gelhar and Axness* [1983], who expressed the heterogeneity-induced solute spreading by means of a longitudinal macrodispersion coefficient D_L^{mac} related to the variance σ_{ff}^2 and correlation length l (see (81)) of the log-hydraulic conductivity field $f(\mathbf{s})$ and the mean, ensemble-averaged, groundwater flow velocity $[[v]]$,

$$D_L^{\mathrm{mac}} \equiv \lim_{t\to\infty} D_{11}^{\mathrm{ens}}(t) \propto \sigma_{ff}^2\, l\, [[v]]. \tag{78}$$

Frequently, large-scale transport is modeled by advective-dispersive transport with local-scale dispersion substituted by the (constant) macrodispersion coefficients. The assumptions underlying this approach are that (1) large-scale transport obeys the same dynamic equation as transport on a local scale and (2) (constant) local-scale transport coefficients can simply be substituted by their (constant) large-scale counterparts given as suitably defined ensemble averages.

[187] However, the $D_{ij}^{\mathrm{ens}}(t)$ are, in general, functions of time or transport distance, as discussed, e.g., in section 6.1 [e.g., *Koch and Brady*, 1987, 1988; *Dagan*, 1988; *Neuman and Zhang*, 1990; *Koch and Shaqfeh*, 1992], i.e.,

$$D_{ij}^{\mathrm{ens}} = D_{ij}^{\mathrm{ens}}(t), \tag{79}$$

which implies that large-scale solute dispersion is, in general, non-Fickian. If $D_{ij}^{\mathrm{ens}}(t)$ does not reach a constant value at large times but increases as a power of time,

$$D_{ij}^{\mathrm{ens}}(t) \propto t^{\varkappa} \tag{80}$$

for $0 < \varkappa < 1$, as derived throughout sections 2–4, one has anomalous transport [e.g., *Bouchaud and Georges*, 1990]. Here we review briefly aspects of non-Fickian dispersion as a consequence of the correlation structure of the log-hydraulic conductivity. We note for comparison, recalling, e.g., section 3.1, the CTRW theory also relates the statistical properties of the hydraulic conductivity, through determination of the statistical distribution of $1/v$, directly to the $\psi(t)$ and the memory function $M(t)$. A related treatment based on analysis of transport in streamlines is given by *Di Donato et al.* [2003].

[188] A criterion for $D_{ij}^{\mathrm{ens}}(t)$ to reach a normal dispersive limit (i.e., a constant long time value) is the existence of a correlation length l of the autocorrelation function of the

log-hydraulic conductivity [*Koch and Brady*, 1988], which is defined as [*Dagan*, 1987]

$$l \equiv \left[C^{ff}(0)\right]^{-1} \int_0^\infty C^{ff}(r)\,\mathrm{d}r \tag{81}$$

with $r = |\mathbf{s}|$. Ensemble dispersion in the flow direction, i.e., $D_{11}^{\mathrm{ens}}(t)$, has been investigated in the literature for long-range correlations [e.g., *Koch and Brady*, 1988; *Cushman*, 1991; *Cushman and Ginn*, 1993; *Dagan*, 1994a; *Rajaram and Gelhar*, 1995; *Di Federico and Zhang*, 1999], i.e., for $C^{ff}(\mathbf{s})$ that behave as

$$C^{ff}(\mathbf{s}) \propto |\mathbf{s}|^{-\gamma} \tag{82}$$

with $0 < \gamma < 1$. For such power law correlation functions one obtains anomalous dispersive behavior, and $D_{11}^{\mathrm{ens}}(t)$ increases with a power of time according to

$$D_{11}^{\mathrm{ens}}(t) \propto t^{1-\gamma} \tag{83}$$

i.e., $\varkappa = 1 - \gamma$ in (80).

[189] For the simple case of purely advective transport and perfect correlation (e.g., a perfectly stratified medium with flow aligned with the strata), $\gamma = 0$ in (83):

$$D_{11}^{\mathrm{ens}}(t) \propto t. \tag{84}$$

This case is known also as the ''racetrack'' model (this is not anomalous transport, however, as it is a superposition of normal transport in each track). In the presence of diffusion, longitudinal ensemble dispersion in a stratified medium behaves as [*Matheron and de Marsily*, 1980]

$$D_{11}^{\mathrm{ens}}(t) \propto t^{1/2}. \tag{85}$$

A discussion of this particular behavior is given by, e.g., *Matheron and de Marsily* [1980], *Koch and Brady* [1988], and *Clincy and Kinzelbach* [2001].

[190] The behavior of the ensemble dispersion coefficient has been addressed frequently in the literature for short-range correlated log-hydraulic conductivity, i.e., with exponential or Gaussian-shaped $C^{ff}(\mathbf{s})$ [e.g., *Dagan*, 1984; *Koch and Brady*, 1987; *Dagan*, 1988; *Naff*, 1990; *Shvidler*, 1993]. For such correlation models the criterion (81) for normal transport in the long time limit is fulfilled, and $D_{11}^{\mathrm{ens}}(t)$ increases from the local dispersion coefficient toward its macroscopic constant long time value (78).

6.2.2. Effective Transport Equation

[191] As outlined in section 6.2.1, the temporal variability of the effective transport coefficients indicates that effective transport, in general, cannot be described by the same transport equation as local-scale transport, with the local-scale parameters substituted by their upscaled counterparts. Transport can be assumed to be Fickian only in a long time limit under certain conditions. As pointed out in section 3, the time required to reach such a limit can be exceedingly long. Thus, for realistic transport times and transport distances, effective solute transport cannot, in general, be described by the local-scale dynamics with upscaled (constant) transport coefficients. This presents a problem also for the averaging approaches discussed in section 6.1.1. In fact, effective transport frameworks have been derived by ensemble averaging the local-scale transport equation, leading to equations for the ensemble-averaged concentration distribution that are nonlocal in space and time.

[192] There are various studies dealing with the derivation of an effective transport equation for transport in a random velocity field. By ensemble averaging of (74), one obtains an equation for the ensemble-averaged concentration distribution $c(\mathbf{s}, t)$ [e.g., *Roberts*, 1961; *Koch and Brady*, 1987; *Cushman*, 1991; *Koch and Shaqfeh*, 1992; *Neuman*, 1993; *Deng et al.*, 1993; *Cushman et al.*, 1995; *Hu et al.*, 1995], which typically has the form

$$\frac{\partial}{\partial t} c(\mathbf{s},t) + [[\mathbf{v}]] \cdot \nabla c(\mathbf{s},t) - \nabla\nabla \int \mathrm{d}^d s' \cdot \int_0^t \left[\mathbf{D}\,\delta(t-t')\,\delta(\mathbf{s}-\mathbf{s}') + \mathcal{D}(\mathbf{s}-\mathbf{s}', t-t')\right] c(\mathbf{s}',t')\,\mathrm{d}t' = 0 \tag{86}$$

where the components $\mathcal{D}_{ij}(\mathbf{s}, t)$ of $\mathcal{D}$ denote memory kernels that integrate the effect of spatial heterogeneity into a spatiotemporal nonlocal effective dispersive flux. In other work, there is also a nonlocal contribution to the velocity term [*Cushman et al.*, 1994; *Cushman and Moroni*, 2001].

[193] Equation (86) represents, in fact, a closure problem as the memory kernel $\mathcal{D}_{ij}(\mathbf{s}, t)$ depends implicitly on the concentration in a single realization. The equation can be closed by using a perturbation expansion for the concentration distribution in the fluctuations of the underlying random fields, which leads to a perturbation expression for the memory function. It turned out that a second-order approximation can lead to inconsistencies for the average spatial concentration profiles [*Koch and Shaqfeh*, 1992; *Neuman*, 1993]. Therefore higher-order closure schemes resulting in partial summations of the perturbation expression for $\mathcal{D}_{ij}(\mathbf{s}, t)$ have been invoked [e.g., *Koch and Shaqfeh*, 1992; *Neuman*, 1993; *Zhang and Neuman*, 1995; *Cushman et al.*, 2002]. *Dentz and Berkowitz* [2005] solved such a closure problem for transport under random adsorption by deriving an exact nonperturbative expression for the memory function in terms of the heterogeneity distribution. The ensemble dispersion coefficient $D_{ij}^{\mathrm{ens}}(t)$ is related to $\mathcal{D}_{ij}(\mathbf{s}, t)$ by [e.g., *Koch and Brady*, 1987, 1988]

$$D_{ij}^{\mathrm{ens}}(t) = D_{ij} + \int \mathrm{d}^d s \int_0^t \mathcal{D}_{ij}(\mathbf{s}, t')\,\mathrm{d}t'. \tag{87}$$

[194] Note that the CTRW formulation, i.e., the GME (2), has no closure problem, because the $\phi(\mathbf{s}, t)$ is based on a physical model of transport (e.g., (18)) and is not a function of $c(\mathbf{s}, t)$. Therefore, as discussed in section 2.2, the general nonlocal equation (2) is not related to the effective transport equations outlined above. In other words, the differences

between nonlocal formulations of transport lie in the kernel of the integrodifferential equation. This kernel is the crucial part and is dependent on the characteristics of a physical model. In the case of the CTRW the starting point is the ensemble average of the kinetics embodied in (1).

6.2.3. Discussion

[195] Most stochastic approaches to subsurface transport model local-scale transport by a spatially and temporally local transport equation that is characterized by spatially and temporally varying transport parameters. Thus nonlocal effects caused by small-scale heterogeneities are not taken into account in the ''traditional'' stochastic approaches.

[196] While not a prerequisite, a local-scale ADE-like equation is almost invariably chosen as the starting point for development of an upscaled transport equation. As discussed in section 6, the application of (74) to describe local-scale transport in a saturated aquifer assumes that transport is Fickian in an homogeneous region. However, in real aquifers the ''homogeneous'' regions are restricted in size. As shown repeatedly in section 3, and pointed out by, e.g., *Levy and Berkowitz* [2003] and *Cortis et al.* [2004a], tracer transport even over small domain sizes is not necessarily Fickian, and flow and transport do not ''homogenize'' on the same temporal and spatial scales. The effect of non-Fickian transport characteristics at the local scale remains to be explored within the framework of the stochastic approach.

[197] A closely related problem arises: In terms of defining a local-scale equation one must ask, How local is ''local''? The heterogeneity of the conductivity field is often used to fix the heterogeneity of the velocity field (through Darcy's law), neither of which is well defined at the pore scale. As such the assumed local transport equation can be considered applicable only at some scale larger than that of the pores. In the same sense the relevant stochastic field is thus not the conductivity field but rather the velocity field.

[198] Discussion of a number of issues related to the very existence and definition of ''macrodispersion'' is in order. First, expressions for the macrodispersion, e.g., the development of (78), are valid only in a large time, Fickian regime for transport. As discussed extensively in section 3, this limit is rarely, if ever, reached in practice. Second, the cutoff at which macrodispersion is reached is not well defined because of memory effects (recall section 3.7). Thus the classical ''macrodispersion'' parameter is not necessarily an ''absolute'' quantity. Another interesting perspective that questions whether macrodispersion coefficients exist is given by *Lowe and Frenkel* [1996].

[199] Furthermore, conceptual questions that remain to be fully addressed refer to the issue of self-averaging [e.g., *Bouchaud and Georges*, 1990; *Clincy and Kinzelbach*, 2001; *Eberhard*, 2004], i.e., determining the length and timescales over which the ensemble average ''observables'' are representative (and not artifacts of the statistical ensemble) of the corresponding observables in a single realization; that is, their sample to sample fluctuations become small. Examples for when this self-averaging property does not apply are transport in a stratified medium [*Clincy and Kinzelbach*, 2001; *Eberhard*, 2004] and transport in a randomly fluctuating transient velocity field in an homogeneous medium [*Dentz and Carrera*, 2003]. *Neuman* [1993] and [*Guadagnini and Neuman*, 2001] developed moment equations for the ensemble moments of the solute concentration to study the uncertainty of the average solute concentration due to the sample-to-sample fluctuations of the solute concentration from realization to realization of the underlying random fields.

[200] Practical application of the stochastic approach to quantify the full evolution of a migrating contaminant plume remains a key issue. The overwhelming emphasis of such studies has focused on spatial moment characterizations of tracer plume migration and/or determination of the macrodispersion parameter. The extent of this limitation is particularly notable when referring to the discussion in section 3.4.2 and the figures referred to therein: It is unfortunate, and somewhat perplexing, that with the intense efforts of the last decades on stochastic analyses, consideration of full, measured BTCs and efforts to fit them has rarely been attempted. Clearly, for real applications we must consider the full spatial and temporal evolution of a migrating tracer plume. A criticism to this effect was voiced by *Gelhar* [1997, p. 174], who stated that the emphasis of the ''stochastic approach'' over the previous decade had been on ''... theoretical refinements of practically insignificant, but conveniently solvable, problems....''

[201] To conclude, both the CTRW approach and the effective transport description resulting from a stochastic transport model describe effective solute transport in terms of nonlocal transport equations. The basic difference between the CTRW approach and the others discussed in this section lies in the basic starting point to quantify transport and in the method to account for the impact of fluctuations present in disordered systems. In each chosen locale the CTRW model incorporates a full pdf of the range of transition rates composing the transport. A good illustration of this difference is given by *Bijeljic and Blunt* [2006], as discussed in section 3.8. The excellent comparison they obtained for the experimentally measured macrodispersion as a function of *Pe* required quantification of the β, derived from the full pdf of the interpore transit times (recall Figures 27–29). This key information is suppressed by the averaging over a locale to obtain coefficients for the ADE. The correct determination of macrodispersion requires full knowledge of the entire time dependence of the spatial moments of the plume (i.e., a solution of a nonlocal-in-time equation) together with the cutoff time t_2 (recall Figures 20–23).

7. ALTERNATIVE EFFECTIVE TRANSPORT FORMULATIONS

[202] Recent attention has been focused on two alternative formulations to treat solute transport, based on the multirate mass transfer (MRMT) approach [*Pfister and Scher*, 1978; *Roth and Jury*, 1993; *Harvey and Gorelick*, 1995; *Haggerty and Gorelick*, 1995; *Haggerty et al.*, 2000;

Carrera et al., 1998] and on the use of fractional differential equations (FDE) [*Benson et al.*, 2000; *Metzler and Klafter*, 2000; *Baeumer et al.*, 2001; *Schumer et al.*, 2003]. We focus below on these MRMT and FDE approaches, placing them into context with CTRW theory.

[203] We note, parenthetically, that two other oft quoted formulations to quantify solute transport are based on transfer functions [*Jury*, 1982; *Jury et al.*, 1986] and stochastic convective [*Dagan and Bresler*, 1979; *Sposito et al.*, 1986] concepts. The transfer function approach is a general descriptive method to characterize solute flux in heterogeneous media; a medium is treated as a black box, and heterogeneity-induced effective transport mechanisms are not specified. Transport of a solute is thus characterized by a transfer function (Green's function), which represents the solute arrival time distribution at a control plane in response to an instantaneous pulse at an input plane. The stochastic convective model is essentially a specific implementation of the transfer function approach, which assumes that solute transport takes place in independent stream tubes with constant flow velocity. There is no mass exchange between stream tubes by local dispersion. The influence of a broad range of flow velocities on solute arrival is then taken into account by averaging the solution of the one-dimensional ADE [*Dagan and Bresler*, 1979; *Sposito et al.*, 1986]. Determination of the number and properties of the stream tubes is ill-defined and nonunique, and the method provides only a purely empirical fit to existing measurements. Another drawback with these approaches is that parameters fit for a specific BTC are not generally relevant for BTC fits or predictions at other times or distances.

7.1. Multirate Mass Transfer

[204] The MRMT approach distinguishes mobile and immobile solute fractions. MRMT thus models the effect of medium heterogeneities on effective solute transport by a distribution of typical mass transfer (exchange) times between mobile and immobile solute phases. The discussion of MRMT includes naturally multiporosity and mobile-immobile models, matrix diffusion, multiple trapping, and transport under linear first-order reactions. MRMT has been applied to quantify transport of linearly sorbing solutes [*Sardin et al.*, 1991; *Roth and Jury*, 1993], and transport in media characterized by regions of fast and very slow solute transport. In the latter context one encounters multiporosity and, as a special case, double porosity models [*Feehley et al.*, 2000] and mobile-immobile [*Nkedi-Kizza et al.*, 1984] and matrix diffusion models [e.g., *Glueckauf*, 1955; *Neretnieks*, 1980; *Rao et al.*, 1980; *Cunningham et al.*, 1997; *Carrera et al.*, 1998]. These are all summarized under the name MRMT models [*Pfister and Scher*, 1978; *Haggerty and Gorelick*, 1995; *Haggerty et al.*, 2000] because they describe linear mass transfer from mobile to immobile solute phases.

[205] In MRMT the total solute concentration $c(\mathbf{s}, t)$ is decomposed into a mobile (subscript m) and an immobile (subscript im) fraction:

$$c(\mathbf{s},t) = c_{\mathrm{m}}(\mathbf{s},t) + c_{\mathrm{im}}(\mathbf{s},t). \tag{88}$$

The temporal change in the total concentration is balanced by the divergence of advective and dispersive flux in the mobile phase, which yields together with (88) the transport equation [*Roth and Jury*, 1993; *Haggerty and Gorelick*, 1995]

$$\frac{\partial c(\mathbf{s},t)}{\partial t} = -\mathbf{v}\cdot\nabla c_{\mathrm{m}}(\mathbf{s},t) + \nabla\cdot\mathbf{D}\,\nabla c_{\mathrm{m}}(\mathbf{s},t). \tag{89}$$

Equation (89) is closed by a linear relation between the mobile and immobile concentrations, such as given by (91) [*Pfister and Scher*, 1978] or by (97) [*Carrera et al.*, 1998; *Haggerty et al.*, 2000; *Dentz and Berkowitz*, 2003].

[206] To better understand this relation between $c_{\mathrm{m}}(\mathbf{s}, t)$ and $c_{\mathrm{im}}(\mathbf{s}, t)$, we first start with a multiple trapping process, which is a specific case of MRMT that involves first-order transitions into and out of immobilizing sites (traps). This was the first example of MRMT shown to be a subset of CTRW [*Schmidlin*, 1977; *Noolandi*, 1977; *Pfister and Scher*, 1978]. In this case we can write

$$\begin{aligned} \tfrac{\partial c_{\mathrm{im}}(\mathbf{s},t)_i}{\partial t} &= c_{\mathrm{m}}(\mathbf{s},t)\omega_i - c_{\mathrm{im}}(\mathbf{s},t)_i W_i \\ c_{\mathrm{im}}(\mathbf{s},t) &= \Sigma_i c_{\mathrm{im}}(\mathbf{s},t)_i, \end{aligned} \tag{90}$$

where ω_i is the capture (trapping) rate and W_i is the release rate of the ith trap level. The traps are characterized by their level (e.g., energy) and not by their spatial position. A configuration average over these positions of the trap sites is assumed. We first solve (90) with a $\mathcal{L}$ for $c_{\mathrm{im}}(\mathbf{s}, t)_i$, sum over i to obtain $\tilde{c}_{\mathrm{im}}(\mathbf{s}, u)$, and then use the $\mathcal{L}$ of (88) to write $\tilde{c}_{\mathrm{m}}(\mathbf{s}, u)$ in terms of $\tilde{c}(\mathbf{s}, u)$,

$$\begin{aligned} \tilde{c}_{\mathrm{im}}(\mathbf{s},u)_i &= \tilde{c}_{\mathrm{m}}(\mathbf{s},u)[\omega_i/(u+W_i)], \\ \tilde{c}_{\mathrm{m}}(\mathbf{s},u) &= \tilde{c}(\mathbf{s},u)\left[1+\sum_i \frac{\omega_i}{u+W_i}\right]^{-1}. \end{aligned} \tag{91}$$

We insert this last relation into the $\mathcal{L}$ of (89) and compare it to (33) to obtain the memory function

$$\tilde{M}(u) = \left[1+\sum_i \frac{\omega_i}{u+W_i}\right]^{-1}, \tag{92}$$

and using (29), we solve for $\tilde{\psi}(u)$ to formally obtain

$$\tilde{\psi}(u) = \left[1+u\bar{t}+u\bar{t}\sum_i \frac{\omega_i}{u+W_i}\right]^{-1}. \tag{93}$$

[207] The ability to cast the MRMT model into the form of (93) demonstrates that MRMT is a special case of CTRW. The relationship only makes sense in the low-u limit, because if the characteristic time is chosen to be the trapping time $\bar{t} = (\Sigma_i \omega_i)^{-1}$ (which is the average time in the mobile state before being trapped again), then the limit $\mu \equiv u\bar{t} \ll 1$ corresponds to t larger than the trapping time, i.e., many trapping events. Hence the random walk is simply a series of many transitions between traps via a mobile state.

[208] There are two time domains in this motion, the time spent in the mobile phase and the range of release times from the traps. The latter dominates the transport. The $u\bar{t} \ll 1$ limit of (93) can be obtained by approximating (93) as

$$\tilde{\psi}(u) = 1 - u\bar{t} - u\bar{t}\sum_i \frac{\omega_i}{u + W_i} + \dots. \quad (94)$$

Neglecting the $u\bar{t}$ term and substituting with $u/(u + W_i) = 1 - W_i/(u + W_i)$, we then have

$$\tilde{\psi}(u) \approx \sum_i \frac{\bar{t}\omega_i W_i}{u + W_i} \quad (95)$$

or

$$\psi(t) \approx \sum_i \chi_i W_i \exp(-W_i t), \quad (96)$$

where $\chi_i = \bar{t}\omega_i$ is the probability to be trapped in the ith level.

[209] One can generalize (91) so that the discrete enumeration of levels by i or the release time W_i^{-1} can be converted into a continuous one $W_i^{-1} \rightarrow t_r$, with a distribution $\mathcal{P}(t_r)$ of "levels," tagged by t_r [e.g., *Carrera et al.*, 1998; *Haggerty et al.*, 2000; *Dentz and Berkowitz*, 2003],

$$c_{\mathrm{im}}(\mathbf{s}, t) = \int_0^\infty \mathcal{P}(t_r)\,\omega(t_r)\,\mathrm{d}t_r \int_0^t \mathcal{R}[(t - t')/t_r]\, c_{\mathrm{m}}(\mathbf{s}, t')\,\mathrm{d}t'. \quad (97)$$

To see the connection with the multiple trapping case, we insert $\mathcal{P}(t_r) = \Sigma_i \delta(t_r - W_i^{-1})$ and $\mathcal{R}(t/t_r) = \exp(-t/t_r)$, the transfer function, into the $\mathcal{L}$ of (97) and define $\omega_i \equiv \omega(W_i^{-1})$ to obtain the $\tilde{c}_{\mathrm{im}}(\mathbf{s}, t)$ in (91).

[210] The particular form of the transfer function $\mathcal{R}(t/t_r)$ and the trapping rate $\omega(t_r)$ depends on the particular mass exchange mechanism between mobile and immobile regions. The function $\mathcal{P}(t_r)$ is, more generally, the distribution of mass transfer times, which reflects the influence of small-scale medium heterogeneities on the effective transport behavior. A "global" transfer function can then be defined by

$$\varphi(t) \equiv \int_0^\infty \mathcal{P}(t_r)\,\omega(t_r)\,\mathcal{R}(t/t_r)\,\mathrm{d}t_r. \quad (98)$$

[211] Explicit functional forms of $\mathcal{P}(t_r)$, $\omega(t_r)$, and $\mathcal{R}(t/t_r)$ can be specified for different physical situations [e.g., *Dentz and Berkowitz*, 2003]. In sections 7.1.1 and 7.1.2 we consider transport involving diffusive and "first-order" mass transfer.

7.1.1. Diffusive Mass Transfer

[212] For diffusive mass transfer the trapping rate $\omega(t_r) = \theta/t_r$, where θ denotes the volume ratio between the mobile and immobile regions. In this case, $\mathcal{P}(t_r)$ denotes the distribution density of typical diffusion times in the immobile regions. For diffusion into spherical immobile regions the transfer function is given by [*Haggerty and Gorelick*, 1995; *Dentz and Berkowitz*, 2003]

$$\mathcal{R}(t/t_r) = t_r\,\mathcal{L}^{-1}\left\{\frac{3}{\sqrt{u\,t_r}}\left[\coth\left(\sqrt{u\,t_r}\right) - \frac{1}{\sqrt{u\,t_r}}\right]\right\}, \quad (99)$$

while for diffusion into layered immobile regions one finds

$$\mathcal{R}(t/t_r) = t_r\,\mathcal{L}^{-1}\left[\frac{1}{\sqrt{u\,t_r}}\tanh\left(\sqrt{u\,t_r}\right)\right]. \quad (100)$$

7.1.2. "First-Order" Mass Transfer

[213] For "first-order" mass transfer between mobile and immobile regions (e.g., the trapping case above [*Pfister and Scher*, 1978; *Haggerty et al.*, 2000]) and for linear kinetic adsorption [*Roth and Jury*, 1993] the transfer function is given by

$$\mathcal{R}(t/t_r) = \exp(-t/t_r). \quad (101)$$

The functional form of $\omega(t_r)$ depends on the particular trapping mechanism. For purely advective trapping, $\omega(t_r) = \omega_o$ is constant; for "diffusive" trapping, i.e., if diffusive mass transfer is to be mimicked by a first-order mechanism, then $\omega(t_r) = \theta/t_r$ [*Dentz and Berkowitz*, 2003].

7.1.3. Synthesis

[214] We now proceed by considering, in the same way as above, the more general case (97). We take the $\mathcal{L}$ of (97) and of (98) to obtain $\tilde{c}_{\mathrm{im}}(\mathbf{s}, u) = \tilde{c}_{\mathrm{m}}(\mathbf{s}, u)\tilde{\varphi}(u)$; combining this expression with the $\mathcal{L}$ of (88), we have $\tilde{c}_{\mathrm{m}}(\mathbf{s}, u) = \tilde{c}(\mathbf{s}, u)/[1 + \tilde{\varphi}(u)]$. Inserting this into the $\mathcal{L}$ of (89), a transport equation for the total concentration $\tilde{c}(\mathbf{s}, u)$, and comparing it to the corresponding transport equation for the decoupled CTRW (33), we find a relation between $\tilde{\varphi}(u)$ and the transition time distribution $\tilde{\psi}(u)$:

$$\tilde{\varphi}(u) = \frac{1 - \tilde{\psi}(u)\,(1 + u\,\bar{t})}{u\,\bar{t}\,\tilde{\psi}(u)}. \quad (102)$$

Recall that here the $\tilde{\psi}(u)$ is based on the multiple trapping process, and it is thus given by, e.g., a function of the form (93) or (95).

[215] It should be emphasized that formally one can define a function $\tilde{\varphi}(u)$ for any well-posed $\tilde{\psi}(u)$ by inserting it into (102). However, an arbitrary $\tilde{\psi}(u)$ will not give rise to a $\tilde{\varphi}(u)$ that makes sense for MRMT. Hence one should develop a physically based $\tilde{\varphi}(u)$ (e.g., in (91)) for a MRMT process and use (102) to find the $\tilde{\psi}(u)$ that corresponds to this particular process (e.g., (93)). The main point is that CTRW encompasses MRMT as a special case. The relation between MRMT and decoupled CTRW on the basis of this expression has been studied in detail by *Pfister and Scher* [1978] and *Dentz and Berkowitz* [2003]. In this review we have considered a number of different $\psi(\mathbf{s}, t)$, e.g., the η function (21), the expression for the RFN (42), and the truncated power law (61). The representation of these functions in a MRMT form, although formally possible, is not a physically meaningful MRMT process.

7.1.4. Generalization

[216] We generalize the above picture to account for a kinetic interaction of particles between immobile states and mobile ones possessing an intrinsic range of transition times (a non-Fickian, dispersive medium). An example of such a picture is the inclusion of matrix diffusion in the analysis of advection-dominated transport in a RFN, discussed in section 3.1. In this picture we couple two different time spectra, one within the dispersive medium due to the heterogeneity and the other due to the distribution of transfer times from the immobile states.

[217] It is straightforward within the CTRW framework to quantify this picture by inserting traps into a system where the motion of the mobile fraction is governed by a $p(\mathbf{s})\psi(t)$. In this situation, instead of using the $\mathcal{L}$ of (89) as the starting equation, we use

$$u\tilde{c}(\mathbf{s},u) - c_0(\mathbf{s}) = -\tilde{M}(u)\left[\mathbf{v}_\psi \cdot \nabla \tilde{c}_m(\mathbf{s},u) - \mathbf{D}_\psi : \nabla\nabla \tilde{c}_m(\mathbf{s},u)\right]. \tag{103}$$

This equation accounts for the overall transport through the (heterogeneous) domain, with the $\psi(t)$ and $\tilde{M}(u)$ being defined accordingly, e.g., as discussed in section 3. In addition, we consider the effect of local traps by retaining the kinetics of the immobile fraction expressed by (90) and (91). Inserting the relation (91) (between $\tilde{c}_m(\mathbf{s}, u)$ and $\tilde{c}(\mathbf{s}, u)$) into (103), we derive an effective memory function, $\tilde{M}(u)_{\text{eff}}$, for this combined motion,

$$\tilde{M}(u)_{\text{eff}} = \tilde{M}(u)/[1 + \Sigma_i \omega_i/(u + W_i)] \tag{104}$$

or, more generally, from (97) and (98),

$$\tilde{M}(u)_{\text{eff}} = \tilde{M}(u)/[1 + \tilde{\varphi}(u)]. \tag{105}$$

[218] Using the definition of a memory function (29), written in terms of $\tilde{\psi}(u)_{\text{eff}}$ as $\tilde{M}(u)_{\text{eff}} = \bar{t}u\tilde{\psi}(u)_{\text{eff}}/[1 - \tilde{\psi}(u)_{\text{eff}}]$, one can formally define an effective $\psi(u)_{\text{eff}}$ corresponding to $\tilde{M}(u)_{\text{eff}}$,

$$\tilde{\psi}(u)_{\text{eff}} = \tilde{\psi}(u)/\{1 + \tilde{\varphi}(u)[1 - \tilde{\psi}(u)]\}. \tag{106}$$

Note that in (106) one cannot substitute $\tilde{\varphi}(u)$ as defined in (102) because there $\tilde{\psi}(u)$ refers to the specialized MRMT case. Here $\tilde{\psi}(u)$ describes the dispersive medium of the mobile states. In the limit $u \to 0$, $\tilde{\psi}(0)_{\text{eff}} = 1$ as long as $\tilde{\varphi}(u)[1 - \tilde{\psi}(u)] \to 0$. The latter condition is an effective limit on how broad the range of mobile times can be with respect to the range of release times.

7.2. Fractional Derivative Equations

[219] FDE formulations to quantify transport have received attention in recent years [e.g., *Benson et al.*, 2000; *Metzler and Klafter*, 2000; *Baeumer et al.*, 2001; *Schumer et al.*, 2003; *Zhang et al.*, 2005]. A detailed treatise is given by *Metzler and Klafter* [2000]. It should be recognized that the term "fractional" can refer to fractional order differentiation in time or space or both. Moreover, a number of definitions for fractional operators exist [*Metzler and Klafter*, 2000]. In principle, one can derive both temporal and spatial FDEs from a limiting form of the CTRW solution (8) by expanding $\Lambda(\mathbf{k}, u)$ in (8) for small values of u and $\mathbf{k}$, rearranging the equation, and using appropriately defined operators. FDEs have also been demonstrated to be special cases of other transport formulations [e.g., *Cushman and Ginn*, 2000]. Here we demonstrate that the "usual" temporal fractional derivative equation for transport is a specialized, asymptotic limit of the CTRW formulation. We also discuss other limitations of using the temporal and spatial FDEs for transport modeling in porous media systems. It is interesting to note that while the (temporal) FDE is usually written phenomenologically, as a "generalized analog" to the ADE, its underlying physical and mathematical picture is elucidated when seen as a limited subset of the CTRW formulation.

[220] The temporal FDE can be written as [*Metzler and Klafter*, 2000]

$$\frac{\partial c(\mathbf{s},t)}{\partial t} = -\frac{\partial^{1-\beta}}{\partial t^{1-\beta}}\{\mathbf{v}'(\mathbf{s}) \cdot \nabla c(\mathbf{s},t) - \nabla \cdot [\mathbf{D}'(\mathbf{s}) \cdot \nabla c(\mathbf{s},t)]\}, \tag{107}$$

with the definition of the operator

$$\frac{\partial^{-\gamma}}{\partial t^{-\gamma}} c(\mathbf{s},t) \equiv \frac{1}{\Gamma(\gamma)} \int_0^t \frac{c(\mathbf{s},t')}{(t-t')^{1-\gamma}} dt'. \tag{108}$$

In (107) the primes on $\mathbf{v}'$ and $\mathbf{D}'$ indicate that these quantities do not have the same dimensions as usual; that is, $\mathbf{v}'$ has dimensions $[L]/[t^\beta]$. Another advantage of the CTRW transport equations, e.g., (33), is that we can define a dimensionless time $\tau \equiv t/\bar{t}$, where $\bar{t}$ is a time unit determined by the physical model for $\psi(t)$ (as in the examples in section 3). The comparison between FDE and CTRW must include this difference in time units. The operator (108) was constructed to possess the important relation of its $\mathcal{L}$,

$$\mathcal{L}[\partial^{-\gamma} c(\mathbf{s},t)/\partial t^{-\gamma}; u] = u^{-\gamma} \tilde{c}(\mathbf{s},u). \tag{109}$$

[221] Commuting the operations in (107) and then performing the $\mathcal{L}$ of (107), we obtain the working transport equation (33) with the special value for $\tilde{M}(u) \propto u^{1-\beta}$. This value of $\tilde{M}(u)$ corresponds to a $\tilde{\psi}(u)$ with the asymptotic form (44), with $u\bar{t} \ll 1$. Hence the temporal FDE is a special case of CTRW restricted to a decoupled, asymptotic power law form for $\psi(t)$ over the range $0 < \beta < 1$.

[222] To obtain a $\tilde{\psi}(u)$ for all u which would be equivalent to (107), we equate (29) to the dimensionless form $u^{1-\beta}(\bar{t}/c_\beta)$, with the result

$$\tilde{\psi}(u) = \frac{1}{1 + c_\beta u^\beta}, \qquad 0 < \beta < 1. \tag{110}$$

The $\mathcal{L}^{-1}$ of (110) is a simple exponential for $\beta = 1$; $\psi(t) \sim t^{-1-\beta}$ for $t/\bar{t} \gg 1$ and is undefined for small $t/\bar{t}$. Hence the solutions of (107) are applicable only in the asymptotic limit and are not necessarily physically meaningful for small $t/\bar{t}$. The time range over which the FDE is actually reliable remains an open question because it depends on the value of

$\bar{t}$, which is not defined naturally in the form of (107). Moreover, because this FDE solution is limited to a fixed β and $0 < \beta < 1$, it cannot account for an evolution to a Fickian regime.

[223] We stress that these properties of the FDE are in sharp contrast to the CTRW models for $\psi(t)$ discussed in, e.g., sections 3.5 and 3.7 that provide the complete temporal solution. Recall also that as shown in section 3, analysis of laboratory and field data reveals many instances for which β falls in the range $1 < \beta < 2$.

[224] One apparent attraction of the FDE to some proponents is its pde form which is similar to the familiar ADE. However, the operator (108) is simply a definition, and actual evaluation of (107) follows methods such as described to solve (33). More significantly, the pde form of the CTRW transport equation (33) contains the FDE (107) as a particular case, and its evaluation for a range of BCs is straightforward as discussed in section 2.6.

[225] In contrast to the temporal FDE a spatial FDE assumes a transition time distribution $\psi(t)$ with a finite first moment and a transition length distribution $p(\mathbf{s})$ with a diverging second moment. This latter condition is unphysical, implying that some particles must execute long jumps instantaneously. This case can be shown to be a Markovian process (rather than a temporally based semi-Markovian one) called a Lévy flight. It is important to recognize that a Lévy flight refers to a random movement in space, where the length of the transitions is considered over discrete steps, but time is not involved. Lévy walks, on the other hand, attach a time "penalty" by assigning a velocity to each transition in space. In the simplest case this velocity is constant; relaxation of this condition leads back to the more general CTRW formulation of section 2.3 [*Klafter et al.*, 1987; *Shlesinger et al.*, 1993]. Lévy walks cannot be described in terms of simple fractional transport equations [*Metzler*, 2000].

[226] Application of a spatial FDE to tracer migration in geological formations demands a domain that contains "streaks" of high and low hydraulic conductivity, arranged so as to lead to particle transitions of high and low velocity. In other words, the physical picture of a Lévy flight requires a wide distribution of streak lengths to obtain a non-Fickian distribution of particle transitions. Recall though that non-Fickian patterns arise even without the clear presence of such a conductivity distribution (e.g., Figure 1). Moreover, even within a long streak, particle scatter will reduce or eliminate the number of long excursions.

[227] The spatial FDE thus uses the power law form $p(\mathbf{s}) \sim |\mathbf{s}|^{-1-\varsigma}$, $0 < \varsigma < 2$, for the transition length, which is the characteristic function of a centered and symmetric Lévy distribution.

[228] For $\varsigma \geq 2$ one recovers the usual Gaussian behavior. With an asymptotic (small $\mathbf{k}$) form for $p(\mathbf{s})$ one can obtain [*Metzler et al.*, 1998] a spatial FDE

$$\frac{\partial}{\partial t} c(\mathbf{s}, t) + \mathbf{v} \cdot \nabla c(\mathbf{s}, t) = D^{\varsigma} \nabla^{\varsigma} c(\mathbf{s}, t), \tag{111}$$

where here D is a generalized diffusion parameter.

[229] In terms of the CTRW framework a power law $p(\mathbf{s})$ can be considered in a decoupled form of $\psi(\mathbf{s}, t)$ or, alternatively, directly in a coupled $\psi(\mathbf{s}, t)$. With regard to the former case we showed in section 3.3 that a power law $p(\mathbf{s})$ is not generally required; an example of the latter case is provided in section 3.1.

[230] We conclude this section by pointing out also that the FDE approach, both temporal and spatial, does not recognize the transport velocity to be fundamentally different from the fluid velocity, as discussed already in sections 3.5 and 6.1.2. Thus the FDE places the mechanism for non-Fickian behavior entirely on the value of the exponent controlling the (spatial or temporal) power law distribution. We stress once again that in contrast the CTRW formulation (33) is more comprehensive: The memory function accounts for the nonlocal-in-time dispersion, whereas $\mathbf{D}_{\psi}$ provides a measure of the local-in-space dispersion.

8. CONCLUDING REMARKS

[231] Quantification of contaminant transport in geological formations has been a long-standing problem. The difficulty in capturing the complexities of tracer plume migration patterns suggests that local, small-scale heterogeneities cannot be neglected; we have shown that these unresolvable heterogeneities contribute significantly to the occurrence of non-Fickian transport. Indeed, BTCs of passive tracers in even macroscopically "homogeneous" granular materials exhibit non-Fickian features: Early and late arrival times are observed to differ systematically from theoretical predictions based on solution of the ADE for uniform porous media. Even in these small-scale, "homogeneous" domains, subtle and residual pore-scale disorder effects can account for these observations.

[232] We have reviewed a recent, different approach to this problem based on a CTRW framework. The theory developed within this framework is structured by a conceptual picture of transport as a sequence of particle transfer rates. The starting point to arrive at the CTRW is the master equation, which describes the kinetics of the probability of site occupancy, incorporating these rates, for a single realization of an heterogeneous medium. The ensemble-averaged ME is the GME, which we show is equivalent to the CTRW, and serves as the transport equation. A particularly convenient approximation of this equation is the pde (33) "similar" in form, in Laplace space, to the well-known ADE. However, this equation is fundamentally different because it incorporates the decisive memory term, $\tilde{M}(u)$, as well as "generalized" $\mathbf{v}_{\psi}$ and $\mathbf{D}_{\psi}$ terms. Equation (33) can also be written in a non-Laplace transformed, nonlocal-in-time form, (35), which can be compared directly to the familiar ADE and solved numerically without requiring Laplace inversion. Boundary value solutions of this transport equation yield tracer migration patterns that are in excellent agreement with the entire series of observations.

[233] On this basis we can state that the CTRW framework represents a powerful and effective means to quantify transport in a wide range of porous and fractured media. It

enables calculation of both BTCs and the full temporal and spatial evolution of contaminant plumes, covering both the premacrodispersion and macrodispersion regime time ranges. Further, as the calculation does not resort to using perturbation theory, the results are valid for strongly heterogeneous formations (e.g., log hydraulic conductivity variance >10). The CTRW theory can be extended naturally to treat transport in nonstationary domains with specific conditioning information.

[234] The CTRW equation is governed by the pdf $\psi(\mathbf{s}, t)$, which characterizes particle transitions in space and time. As demonstrated throughout sections 2, 3, and 4, the choice of the functional form of $\psi(\mathbf{s}, t)$ for any given application must be dictated by the physics of the flow field, which are determined by the particular heterogeneous system. A generic form of $\psi(\mathbf{s}, t)$ can be seen in Figure 3. It is calculated from (18) and (19), the ensemble average of a functional of the rates w appearing in the ME. The type of time dependence and parameter values of $\psi(\mathbf{s}, t)$ determine the nature of the transport. For example, if the duration of observations coincides with a slow power law type of t dependence, then the transport will be strongly influenced by (rare) low-velocity regions; as a consequence one need not distinguish explicitly between the contributions of slow advection and diffusion. From a practical, field-scale point of view it is in any case often difficult to make such distinctions. When such distinctions can be made, we have generalized the CTRW transport theory in dispersive media to specifically include interactions with immobile states or matrix diffusion. The direct relationship of $\psi(\mathbf{s}, t)$ to the flow field thus connects it to the fully correlated hydraulic conductivity field. If it is known, the velocity histogram determines $\psi(\mathbf{s}, t)$ (e.g., see section 3.1). What remains is to investigate how best to use known heterogeneity details of a domain to define a specific $\psi(\mathbf{s}, t)$. The transport problem has been successively reduced to this key aspect of the physical modeling.

[235] Issues of parameter fitting and model upscaling arise naturally. Clearly, application of any transport theory requires understanding of the nature of the model parameters in the context of a particular problem of interest. Thus in the context of CTRW, for example, β in the asymptotic $\psi(t)$ function, as used in (47), changes with increases in the flow velocity v. The change in v shifts the time range to different features of the $\psi(t)$ (or $\psi(\mathbf{s}, t)$) function. It is the $\psi(t)$ function that is intrinsic to the physical modeling and not β, which characterizes a feature of $\psi(t)$. In the CTRW framework, as in any other fully descriptive transport modeling approach, fitting parameters cannot, by definition, be determined a priori or solely on the basis of consideration of the hydraulic conductivity distribution. Domain-specific measurements must be used to constrain the model.

[236] As a consequence, and because hierarchies of heterogeneity scales, which cannot be resolved, exist over a broad range of length scales, we must ultimately work at the length scale of interest and not expect to simply "scale up" from, for example, core scale measurements to predictions at the 100 m scale. Thus it is not reasonable to expect that model parameters (whether evaluated for CTRW, the ADE or any other model) determined from, e.g., a laboratory-scale analysis of a 10 cm long rock core sample can necessarily be "scaled up" to realistically capture field-scale behaviors. In other words, we must reconcile with the borders of predictability and accept the reality that model parameters are not scale-independent. We have introduced the hybrid approach to deal with this problem, i.e., to restrict the parameters of $\psi(t)$ to a smaller scale while the larger scale is conditioned by observed trends of a permeability field.

[237] Motivated by these considerations, we have examined underlying assumptions of upscaling formulations and the stochastic approach and their differences with respect to CTRW. With regard to other approaches we presented several analyses that demonstrated the ADE, fractional derivative, and multirate mass transport and mobile-immobile models to be, in fact, specialized subsets within the CTRW framework.

[238] We have suggested that consideration of the temporal-spatial correlations in tracer transport within the CTRW framework represents a change in paradigm. The emphasis of the CTRW approach on temporal aspects of particle transport, induced by spatial heterogeneity, is a key feature. In other words, CTRW theory recognizes that particles experience a distribution of times not just lengths of transitions, as they move through a porous/fractured medium. Importantly, the temporal and spatial scales of homogenization for transport are different than they are for flow. The shift in paradigm is due to the different approach to averaging: a full pdf of local rates versus a local average rate (such as for the ADE and variant formulations). As an example the parameters β and t_2 derived from the pore-scale pdf (e.g., Figure 27) were needed to determine the scaling with Pe of the measured macrodispersion.

[239] We conclude by stating that because natural heterogeneity in geological formations occurs over a broad range of scales, we must recognize the ubiquity of "non-Fickian" transport, which should be considered the expected phenomenon, with Fickian transport being the exception.

APPENDIX A: RANDOM WALKS

[240] The familiar random walk (RW) is a sequence of displacements, each one labeled by an integer step number n. The displacement at each step is governed by a distribution $p(\mathbf{s})$. A series of $\mathbf{s}$ chosen at random from $p(\mathbf{s})$ generates one such RW. The outcome of this procedure executed repeatedly from the same origin can be described by the linear recursion relation

$$P_n(\mathbf{s}) = \sum_{\mathbf{s}'} p(\mathbf{s} - \mathbf{s}') P_{n-1}(\mathbf{s}') \tag{A1}$$

$$\sum_{\mathbf{s}} p(\mathbf{s}) = 1, \tag{A2}$$

where $P_n(\mathbf{s})$ is the probability for the walker to be found at $\mathbf{s}$ after n steps, which in (A1) is determined by a single step

from where it was at $n-1$ steps. For $n \gg 1$ one can assume n to be a continuous variable ($\propto t$) and expand $P_{n-1}(\mathbf{s}')$ in n and $\mathbf{s}-\mathbf{s}'$. The expansion

$$\begin{aligned} P_{n-1}(\mathbf{s}') &\approx P_n(\mathbf{s}') - \frac{\partial}{\partial n}P_n(\mathbf{s}') \\ &\approx P_n(\mathbf{s}) + (\mathbf{s}'-\mathbf{s})\cdot\nabla P_n(\mathbf{s}) \\ &\quad + \frac{1}{2}(\mathbf{s}'-\mathbf{s})(\mathbf{s}'-\mathbf{s}) : \nabla\nabla P_n(\mathbf{s}) - \frac{\partial}{\partial n}P_n(\mathbf{s}) \end{aligned} \tag{A3}$$

is substituted into (A1), with the dyadic symbol (colon) denoting a tensor product.

[241] The lowest-order term is retained on the right side of (A3) in the expansion of $\partial P_n(\mathbf{s}')/\partial n$. Assuming the isotropic case, the second term on the right side of (A3) does not contribute (the first moment of $p(\mathbf{s})$ vanishes). The time variable is introduced by multiplying the resulting equation by r, the number of steps per time (i.e., $n = rt$) to obtain the classic diffusion equation

$$\frac{\partial}{\partial t}P(\mathbf{s},t) = D\nabla^2 P(\mathbf{s},t), \tag{A4}$$

where

$$D \equiv \frac{r}{2}\int d\mathbf{s}\frac{1}{3}s^2 p(\mathbf{s}). \tag{A5}$$

Hence a simple random walk, after many steps, is the same as a diffusion process. There is effectively one rate constant r in this process.

[242] To account for systems with a multitude of rates, the RW is generalized as

$$R_n(\mathbf{s},t) = \sum_{\mathbf{s}'}\int_0^t \psi(\mathbf{s}-\mathbf{s}', t-t')R_{n-1}(\mathbf{s}',t')dt', \tag{A6}$$

where $R_n(\mathbf{s}, t)$ is the probability per time for a walker to just arrive at site $\mathbf{s}$ at time t in n steps and $\psi(\mathbf{s}, t)$ is the probability rate for a displacement $\mathbf{s}$ with a difference of arrival times of t. Clearly, the joint distribution $\psi(\mathbf{s}, t)$ is the generalization of $p(\mathbf{s})$ in (A1). The steps can each now take place at different times. The initial condition is $R_0(\mathbf{s}, t) = \delta_{\mathbf{s},0}\delta(t-0^+)$. The time integral in (A6) sums over all the possible times for the single step to occur. If we are not interested in the number of steps but only in the time to reach $\mathbf{s}$ from the origin, then one can sum over all possible steps and define

$$R(\mathbf{s},t) \equiv \sum_{n=0}^{\infty} R_n(\mathbf{s},t), \tag{A7}$$

and summing over n in (A6), one arrives at the form for the CTRW

$$R(\mathbf{s},t) - \sum_{\mathbf{s}'}\int_0^t \psi(\mathbf{s}-\mathbf{s}', t-t')R(\mathbf{s}',t')dt' = \delta_{\mathbf{s},0}\delta(t-0^+) \tag{A8}$$

as shown in (3).

[243] The equivalence between CTRW (3) and the GME (2) is derived by starting with the $\mathcal{L}$ of (A8)

$$\tilde{R}(\mathbf{s},u) - \sum_{\mathbf{s}'}\tilde{\psi}(\mathbf{s}-\mathbf{s}',u)\tilde{R}(\mathbf{s}',u) = \delta_{\mathbf{s},0} \tag{A9}$$

and substituting (from the $\mathcal{L}$ of (4) and (5) with P instead of c)

$$\tilde{R}(\mathbf{s},u) = u\tilde{P}(\mathbf{s},u)/\left[1-\tilde{\psi}(u)\right]. \tag{A10}$$

The final form is arranged by subtracting $\tilde{\psi}(u)u\tilde{P}(\mathbf{s}, u)/[1-\tilde{\psi}(u)]$ from both sides of (A9) to obtain

$$\begin{aligned} u\tilde{P}(\mathbf{s},u) - \delta_{\mathbf{s},0} = &-\sum_{\mathbf{s}'}\frac{u\tilde{\psi}(\mathbf{s}'-\mathbf{s},u)\tilde{P}(\mathbf{s},u)}{1-\tilde{\psi}(u)} \\ &+\sum_{\mathbf{s}'}\frac{u\tilde{\psi}(\mathbf{s}-\mathbf{s}',u)\tilde{P}(\mathbf{s}',u)}{1-\tilde{\psi}(u)}. \end{aligned} \tag{A11}$$

Now the Laplace transform is inverted to obtain the GME

$$\begin{aligned} \frac{\partial P(\mathbf{s},t)}{\partial t} = &-\sum_{\mathbf{s}'}\int_0^t \phi(\mathbf{s}'-\mathbf{s}, t-t')P(\mathbf{s},t')dt' \\ &+\sum_{\mathbf{s}'}\int_0^t \phi(\mathbf{s}-\mathbf{s}', t-t')P(\mathbf{s}',t')dt' \end{aligned} \tag{A12}$$

with

$$\tilde{\phi}(\mathbf{s},u) = u\tilde{\psi}(\mathbf{s},u)/\left[1-\tilde{\psi}(u)\right] \tag{A13}$$

as shown in (7), the $\mathcal{L}$ of $\phi(\mathbf{s}, t)$.

[244] As discussed in section 2.5, using a form of $\psi(\mathbf{s}, t)$ containing a single rate

$$\psi(\mathbf{s},t) = w(\mathbf{s})\exp(-Wt), \qquad W \equiv \Sigma_{\mathbf{s}}w(\mathbf{s}) \tag{A14}$$

reduces the GME (A12) to a local-in-time equation. It is straightforward to show this result by inserting the $\mathcal{L}$ of (A14)

$$\tilde{\psi}(\mathbf{s},u) = \frac{w(\mathbf{s})}{u+W} \tag{A15}$$

into (A13) to determine $\tilde{\phi}(\mathbf{s}, u) = w(\mathbf{s})$, which is independent of u. Hence $\phi(\mathbf{s}, t) = w(\mathbf{s})\delta(t)$; that is, there is no memory function. This result is less obvious from the CTRW form of the equation (A8) where the single rate form of $\psi(\mathbf{s}, t)$ should reduce (A8) to a Markovian equation. Again, the $\mathcal{L}$ is used; (A15) and (A10) are inserted into (A9) and algebraically rearranged to yield

$$(u+W)\tilde{P}(\mathbf{s},u) - \delta_{\mathbf{s},0} = \sum_{\mathbf{s}'}w(\mathbf{s}-\mathbf{s}')\tilde{P}(\mathbf{s}',u), \tag{A16}$$

which can now be $\mathcal{L}$ inverted to obtain

$$t_\sigma\frac{\partial}{\partial t}P(\mathbf{s},t) + P(\mathbf{s},t) = \sum_{\mathbf{s}'}p(\mathbf{s}-\mathbf{s}')P(\mathbf{s}',t), \tag{A17}$$

where $t_\sigma \equiv 1/W$ and $p(\mathbf{s}) \equiv w(\mathbf{s})/W$. The left side of (A17) is the low-order expansion of $P(\mathbf{s}, t + t_\sigma)$. Hence the Markovian process is

$$P(\mathbf{s}, t + t_\sigma) = \sum_{\mathbf{s}'} p(\mathbf{s} - \mathbf{s}')P(\mathbf{s}', t), \tag{A18}$$

which recursively relates $P(\mathbf{s}, t)$ with discrete time steps of t_σ as in (A1) (i.e., $t/t_\sigma = n$).

[245] On a brief historical note the year 2005 marked the 100 year anniversary of A. Einstein's famous paper on Brownian motion [*Einstein*, 1905]. His analysis involved a recursion relation and an expansion leading to the diffusion equation similar to the development outlined above. His solution of the latter showed for the first time that the displacement (rms) σ of the Brownian particle is proportional to $\sqrt{t}$ and not equal to vt as was being assumed in those years. By using the wrong (latter) dependence, the experimentalists were puzzled by the need to invoke a time-dependent velocity v. Einstein investigated systems with one effective rate (e.g., r above). The different time dependence of the particle displacement ($\sqrt{t}$) is due to the nature of diffusive motion, a result now of common understanding. Seventy years later, observations of the transit time τ_r of electrons in disordered semiconductors presented a similar puzzle. The electron mobility (velocity per unit electric field), considered to be an intrinsic property of the material, was found to depend on sample length, electric field, etc. (any variable that changes the duration of the experiment). Again, the problem was traced to using the wrong relationship for the mean displacement $\bar{\ell}$ of the electron packet. Instead of $\bar{\ell} = vt$ it was discovered that $\bar{\ell} \propto t^\beta$ (equation (46)) [*Scher and Montroll*, 1975]. In this case, however, the difference is ascribed to a wide distribution of rates due to the disorder of the system as we have been discussing in this paper in the geological context.

APPENDIX B: DERIVATION OF RELATION BETWEEN $\psi(s, t)$ AND $w(s', s)$

[246] We cast the master equation into the form of a random walk equation and derive the form for $\psi_{\mathbf{s}',\mathbf{s}}(t)$:

$$\frac{\partial C(\mathbf{s}, t)}{\partial t} = -\sum_{\mathbf{s}'} w(\mathbf{s}', \mathbf{s})C(\mathbf{s}, t) + \sum_{\mathbf{s}'} w(\mathbf{s}, \mathbf{s}')C(\mathbf{s}', t), \tag{B1}$$

where $w(\mathbf{s}', \mathbf{s})$ is the transition rate from $\mathbf{s}'$ to $\mathbf{s}$; the dimension of $\Sigma_{\mathbf{s}} w$ is reciprocal time. The $\psi_{\mathbf{s}',\mathbf{s}}(t)$ plays the role of $\psi(\mathbf{s}, t)$, but it is dependent on the location; that is, each neighborhood is different in a specific representation. Each transition between lattice sites is assigned a transition rate.

[247] Recalling section 2.3, the general equation for the random walker in this system, which is nonstationary in space and stationary in time, is

$$R(\mathbf{s}, t) = \sum_{\mathbf{s}'} \int_0^t \psi_{\mathbf{s},\mathbf{s}'}(t - t')R(\mathbf{s}', t')dt', \tag{B2}$$

where the $\psi_{\mathbf{s},\mathbf{s}'}(t)$ now is an explicit function of position. The correspondence between the ME and the random walk is

$$C(\mathbf{s}, t) = \int_0^t \Psi_{\mathbf{s}}(t - t')R(\mathbf{s}, t')dt', \tag{B3}$$

$$\Psi_{\mathbf{s}}(t) = 1 - \int_0^t \psi_{\mathbf{s}}(t')dt', \qquad \psi_{\mathbf{s}}(t) \equiv \sum_{\mathbf{s}'} \psi_{\mathbf{s}',\mathbf{s}}(t). \tag{B4}$$

The u dependence of the Laplace transform is understood in the following:

$$\tilde{\Psi}_{\mathbf{s}} = \frac{1 - \tilde{\psi}_{\mathbf{s}}}{u} \tag{B5}$$

$$\tilde{C}_{\mathbf{s}} = \tilde{\Psi}_{\mathbf{s}}\tilde{R}_{\mathbf{s}} \quad \longrightarrow \quad \tilde{R}_{\mathbf{s}} = \frac{u\tilde{C}_{\mathbf{s}}}{1 - \tilde{\psi}_{\mathbf{s}}}. \tag{B6}$$

From (B2),

$$\tilde{R}_{\mathbf{s}} - \sum_{\mathbf{s}'} \tilde{\psi}_{\mathbf{s},\mathbf{s}'}\tilde{R}_{\mathbf{s}'} = \delta_{\mathbf{s},0}, \tag{B7}$$

and substituting (B6), we obtain

$$\frac{u\tilde{C}_{\mathbf{s}}}{1 - \tilde{\psi}_{\mathbf{s}}} = \sum_{\mathbf{s}'} \tilde{\psi}_{\mathbf{s},\mathbf{s}'} \frac{u\tilde{C}_{\mathbf{s}'}}{1 - \tilde{\psi}_{\mathbf{s}'}} + \delta_{\mathbf{s},0}. \tag{B8}$$

[248] We now subtract the term $u\tilde{\psi}_{\mathbf{s}}\tilde{C}_{\mathbf{s}}/(1 - \tilde{\psi}_{\mathbf{s}})$ from both sides of (B8) to obtain

$$\frac{u\tilde{C}_{\mathbf{s}}}{1 - \tilde{\psi}_{\mathbf{s}}} - \frac{u\tilde{\psi}_{\mathbf{s}}\tilde{C}_{\mathbf{s}}}{1 - \tilde{\psi}_{\mathbf{s}}} = -\sum_{\mathbf{s}'} \frac{u\tilde{\psi}_{\mathbf{s}',\mathbf{s}}\tilde{C}_{\mathbf{s}}}{1 - \tilde{\psi}_{\mathbf{s}}} + \sum_{\mathbf{s}'} \tilde{\psi}_{\mathbf{s},\mathbf{s}'} \frac{u\tilde{C}_{\mathbf{s}'}}{1 - \tilde{\psi}_{\mathbf{s}'}} + \delta_{\mathbf{s},0}, \tag{B9}$$

which is equal to

$$u\tilde{C}_{\mathbf{s}} - \delta_{\mathbf{s},0} = -\sum_{\mathbf{s}'} \frac{u\tilde{\psi}_{\mathbf{s}',\mathbf{s}}\tilde{C}_{\mathbf{s}}}{1 - \tilde{\psi}_{\mathbf{s}}} + \sum_{\mathbf{s}'} \frac{u\tilde{\psi}_{\mathbf{s},\mathbf{s}'}\tilde{C}_{\mathbf{s}'}}{1 - \tilde{\psi}_{\mathbf{s}'}}. \tag{B10}$$

[249] One now compares the two forms of the $\mathcal{L}$ of the ME, i.e., the $\mathcal{L}$ of (B1) and (B10), to obtain

$$\frac{u\tilde{\psi}_{\mathbf{s}',\mathbf{s}}}{1 - \tilde{\psi}_{\mathbf{s}}} = w(\mathbf{s}', \mathbf{s}). \tag{B11}$$

First, we sum over $\mathbf{s}'$ and solve for $\tilde{\psi}_{\mathbf{s}}$

$$\frac{u\tilde{\psi}_{\mathbf{s}}}{1 - \tilde{\psi}_{\mathbf{s}}} = \sum_{\mathbf{s}'} w(\mathbf{s}', \mathbf{s}), \tag{B12}$$

so that

$$\tilde{\psi}_{\mathbf{s}} = \frac{\sum_{\mathbf{s}'} w(\mathbf{s}', \mathbf{s})}{u + \sum_{\mathbf{s}'} w(\mathbf{s}', \mathbf{s})}. \tag{B13}$$

Substituting this expression for $\tilde{\psi}_{\mathbf{s}}$ into (B11) yields

$$\tilde{\psi}_{\mathbf{s}',\mathbf{s}} = \frac{w(\mathbf{s}',\mathbf{s})}{u + \sum_{\mathbf{s}''} w(\mathbf{s}'',\mathbf{s})}. \tag{B14}$$

Inverting the $\mathcal{L}$, we obtain the result

$$\psi_{\mathbf{s}',\mathbf{s}}(t) = w(\mathbf{s}',\mathbf{s}) \exp\left[-t \sum_{\mathbf{s}''} w(\mathbf{s}'',\mathbf{s})\right]. \tag{B15}$$

APPENDIX C: DEFINITION OF THE DYADIC

[250] For the sake of clarity the dyadic is defined as follows. Noting that

$$\mathbf{ss} = \begin{bmatrix} s_x\mathbf{i}, & s_y\mathbf{j}, & s_z\mathbf{k} \end{bmatrix} \begin{bmatrix} s_x\mathbf{i} \\ s_y\mathbf{j} \\ s_z\mathbf{k} \end{bmatrix} = \begin{bmatrix} s_x s_x\mathbf{ii} & s_x s_y\mathbf{ij} & s_x s_z\mathbf{ik} \\ s_y s_x\mathbf{ji} & s_y s_y\mathbf{jj} & s_y s_z\mathbf{jk} \\ s_z s_x\mathbf{ki} & s_z s_y\mathbf{kj} & s_z s_z\mathbf{kk} \end{bmatrix} \tag{C1}$$

and

$$\nabla\nabla = \begin{bmatrix} \partial_x\mathbf{i}, & \partial_y\mathbf{j}, & \partial_z\mathbf{k} \end{bmatrix} \begin{bmatrix} \partial_x\mathbf{i} \\ \partial_y\mathbf{j} \\ \partial_z\mathbf{k} \end{bmatrix} = \begin{bmatrix} \partial_x\partial_x\mathbf{ii} & \partial_x\partial_y\mathbf{ij} & \partial_x\partial_z\mathbf{ik} \\ \partial_y\partial_x\mathbf{ji} & \partial_y\partial_y\mathbf{jj} & \partial_y\partial_z\mathbf{jk} \\ \partial_z\partial_x\mathbf{ki} & \partial_z\partial_y\mathbf{kj} & \partial_z\partial_z\mathbf{kk} \end{bmatrix}, \tag{C2}$$

then it follows that

$$\begin{aligned} \mathbf{ss} : \nabla\nabla P_n(x,y,z) = & \begin{bmatrix} s_x s_x\mathbf{ii} & s_x s_y\mathbf{ij} & s_x s_z\mathbf{ik} \\ s_j s_x\mathbf{ji} & s_y s_y\mathbf{jj} & s_y s_z\mathbf{jk} \\ s_z s_x\mathbf{ki} & s_z s_y\mathbf{kj} & s_z s_z\mathbf{kk} \end{bmatrix} \\ & \cdot \begin{bmatrix} \partial_x\partial_x\mathbf{ii} & \partial_x\partial_y\mathbf{ij} & \partial_x\partial_z\mathbf{ik} \\ \partial_y\partial_x\mathbf{ji} & \partial_y\partial_y\mathbf{jj} & \partial_y\partial_z\mathbf{jk} \\ \partial_z\partial_x\mathbf{ki} & \partial_z\partial_y\mathbf{kj} & \partial_z\partial_z\mathbf{kk} \end{bmatrix} P_n(x,y,z) \\ & = s_x^2\partial_x^2 P_n(x,y,z) + s_x s_y\partial_x\partial_y P_n(x,y,z) \\ & + s_x s_z\partial_x\partial_z P_n(x,y,z) + s_y s_x\partial_y\partial_x P_n(x,y,z) \\ & + s_y^2\partial_y^2 P_n(x,y,z) + s_y s_z\partial_y\partial_z P_n(x,y,z) \\ & + s_z s_x\partial_z\partial_x P_n(x,y,z) + s_z s_y\partial_z\partial_y P_n(x,y,z) \\ & + s_z^2\partial_z^2 P_n(x,y,z). \end{aligned} \tag{C3}$$

APPENDIX D: LOW u EXPANSION OF $\tilde{\psi}(\mathbf{s}, u)$ FOR THE RANDOM FRACTURE MODEL

[251] We shall carry out the low u (or more precisely the dimensionless form μ) expansion of $\tilde{\psi}(\mathbf{s}, u)$ in (42) and evaluate the functions $p_1(\mathbf{s})$ and $p_2(\mathbf{s})$ in (48) for this case. We rewrite (42) in the modified notation

$$\begin{aligned} \tilde{\psi}(\mathbf{s},u) &= \mathcal{N}(\varrho)\,\cos^{2\beta}\left(\frac{\theta}{2}\right) z^\beta K_\beta(2z), \\ \mathcal{N}(\varrho) &\equiv C_n'\varrho^{3/2}\exp(-\varrho), \\ z &\equiv \sqrt{2\varrho\mu}/\cos\frac{\theta}{2}. \end{aligned} \tag{D1}$$

We expand $K_\beta(2z)$ in an ascending power series in z,

$$K_\beta(2z) = \frac{\pi}{2\sin(\beta\pi)}\left[I_{-\beta}(2z) - I_\beta(2z)\right], \tag{D2}$$

where

$$\begin{aligned} I_{-\beta}(2z) - I_\beta(2z) = {} & z^{-\beta}\sum_{k=0}^{\infty}\frac{z^{2k}}{k!\Gamma(1+k-\beta)} \\ & - z^\beta\sum_{k=0}^{\infty}\frac{z^{2k}}{k!\Gamma(1+k+\beta)} \end{aligned} \tag{D3}$$

$$\begin{aligned} z^\beta K_\beta(2z) = {} & \frac{\pi}{2\sin(\beta\pi)}\left[\sum_{k=0}^{\infty}\frac{z^{2k}}{k!\Gamma(1+k-\beta)}\right. \\ & \left. -z^{2\beta}\sum_{k=0}^{\infty}\frac{z^{2k}}{k!\Gamma(1+k+\beta)}\right]. \end{aligned} \tag{D4}$$

To lowest order in μ (for $z < 1$), (D4) is

$$z^\beta K_\beta(2z) \sim \frac{1}{2}\Gamma(\beta) - \frac{\pi}{2\sin(\beta\pi)}\left[\frac{2\varrho}{\cos^2\left(\frac{\theta}{2}\right)}\right]^\beta \frac{\mu^\beta}{\Gamma(1+\beta)} + \mathcal{O}(\mu), \tag{D5}$$

which when inserted into (D1) is readily seen to be of the form in (48) with

$$\begin{aligned} p_1(\mathbf{s}) &= \mathcal{N}(\varrho)\,\cos^{2\beta}\left(\frac{\theta}{2}\right)\frac{1}{2}\Gamma(\beta) \\ p_2(\mathbf{s}) &= \mathcal{N}(\varrho)\frac{\pi(2\varrho)^\beta}{2\sin(\beta\pi)\Gamma(1+\beta)}. \end{aligned} \tag{D6}$$

An interesting feature of the μ expansion in (D4) is the absence of powers of μ^β. The expansion is of the form of a sum of an entire function and μ^β multiplied by an entire function.

APPENDIX E: ASYMPTOTIC EXPRESSION FOR A POWER LAW $\psi(t)$

[252] We derive, for the sake of completeness, the well-known asymptotic expression for $\tilde{\psi}(u)$ (44) for ψ functions that have a power law tail

$$\psi(t) \sim t^{-1-\beta} \tag{E1}$$

with $0 < \beta < 1$. We start by adding and subtracting 1 from the definition of $\tilde{\psi}(u)$ in the form

$$\tilde{\psi}(u) = 1 - \int_0^\infty (1 - e^{-ut})\psi(t)\,dt. \quad \text{(E2)}$$

We assume that $\psi(t)$ is well approximated by the power law in (E1) beyond some value $t = a$. Hence we can write (E2)

$$\tilde{\psi}(u) = 1 - \int_0^a (1 - e^{-ut})\psi(t)dt - c_1 \int_a^\infty \frac{1 - e^{-ut}}{t^{1+\beta}}dt, \quad \text{(E3)}$$

where c_1 is a constant. The integration variable t in the left integral of (E3) is bounded between two finite values, and hence we can expand the exponential term in this integral and (below) retain the lowest-order term of u:

$$\int_0^a (1 - e^{-ut})\psi(t)\,dt = -\sum_{n=1}^\infty \frac{(-u)^n \bar{t}_a^n}{n!}, \quad \text{(E4)}$$

where

$$\bar{t}_a^n = \int_0^a t^n \psi(t)dt. \quad \text{(E5)}$$

Hence this integral results in terms of order u and higher. We recast the right integral of (E3) with an integration by parts,

$$\int_a^\infty \frac{1 - e^{-ut}}{t^{1+\beta}}dt = -\frac{1 - e^{-ut}}{\beta t^\beta}\Big|_a^\infty + u\int_a^\infty \frac{e^{-ut}}{\beta t^\beta}dt. \quad \text{(E6)}$$

The left term in (E6) is of order u and higher. In the right integral in (E6) we make the substitution $x = ut$, so that

$$u\int_a^\infty \frac{e^{-ut}}{\beta t^\beta}dt = \frac{u^\beta}{\beta}\int_{ua}^\infty x^{-\beta}e^{-x}dx = \frac{u^\beta}{\beta}[\Gamma(1-\beta) - \gamma(1-\beta, ua)], \quad \text{(E7)}$$

where $\gamma(b, x)$ is the incomplete gamma function [*Abramowitz and Stegun*, 1970]. Combining this solution with the integral in (E4) and substituting into (E3) leads to the asymptotic approximation for small u:

$$\tilde{\psi}(u) = 1 - c_2 u^\beta + \frac{a^\beta}{\beta}u\sum_{n=0}^\infty \frac{(-ua)^n}{(1-\beta+n)n!} + \mathcal{O}(u) + \ldots, \quad \text{(E8)}$$

where $c_2 = \Gamma(1 - \beta)c_1/\beta$ is a constant. Note (E8) contains a u^β term plus entire functions of u. This is in contrast to (D4) where the asymptotic expression contains an entire function and μ^β multiplied by an entire function (which is due to the exponential factor in (41)).

NOTATION

Note that only the principal symbols used throughout this report are given in the notation list.

c_β	constant (equation (44)).
$\bar{c}_\beta$	constant (equation (48)).
$c(\mathbf{s}, t)$	ensemble-averaged, normalized concentration.
$C(\mathbf{s}, t)$	normalized concentration.
$\tilde{c}(\mathbf{s}, u)$	Laplace transform of $c(\mathbf{s}, t)$.
$\mathcal{C}(\mathbf{k}, u)$	Fourier transform of $\tilde{c}(\mathbf{s}, u)$.
$c_m(\mathbf{s}, t)$	mobile phase.
$c_{im}(\mathbf{s}, t)$	immobile phase.
$\mathbf{D}$	dispersion tensor.
$\mathbf{D}_\psi$	$\frac{1}{2}\overline{\mathbf{ss}}/\bar{t} = \frac{1}{2}\lvert\mathbf{v}_\psi\rvert\overline{\mathbf{ss}}/\lvert\bar{\mathbf{s}}\rvert \equiv \frac{1}{2}\lvert\mathbf{v}_\psi\rvert\alpha_\psi$.
D^L	longitudinal dispersion.
D^T	transverse dispersion.
$F(\mathbf{s}, t)$	first-passage time distribution.
$f_B(t)$	$\equiv \Sigma_{s2,s3}F(s_1 = L, s_2, s_3, t)$, form of BTC.
$\mathcal{F}$	Fourier transform.
$h(t)$	effective traveltime distribution, see (53).
j	average mass flux.
$K_\beta(z)$	modified Bessel function.
$\mathbf{k}$	Fourier variable (dimensional).
$\mathcal{L}$	Laplace transform.
L	length.
ℓ	length scale.
$\bar{\ell}(t)$	mean plume displacement.
$M(t)$	memory function.
$\tilde{M}(u)$	Laplace transform of $M(t)$.
$p(\mathbf{s})$	probability distribution of transition displacements.
Pe	Peclet number.
$\mathbf{q}$	specific discharge.
$R(\mathbf{s}, t)$	probability per time for particle to arrive at site $\mathbf{s}$ at time t.
$\mathbf{s}, \mathbf{x}$	location in space.
$\bar{\mathbf{s}}$	mean distance, first moment of $p(\mathbf{s})$.
t	time.
$\hat{t}$	first temporal moment (mean time) of $\psi(t)$.
$\bar{t}$	characteristic time.
t_1	median transition time in ψ.
t_2	cutoff time in ψ.
u	Laplace variable (dimensional).
v	velocity (one-dimensional).
$\mathbf{v}$	velocity.
$\mathbf{v}_\psi$	$\bar{\mathbf{s}}/\bar{t}$.
$w(\mathbf{s}, \mathbf{s}')$	particle transition from $\mathbf{s}'$ to $\mathbf{s}$.
$\{w\}$	aggregate of transition rates.
W	total rate, defined as $W = \sum_{\mathbf{s}'} w(\mathbf{s}', \mathbf{s})$.
x, y	large, small length scales.
α	dispersivity.
β	exponent (equation (44)).
γ	exponent (equation (48)).
η	see (52), also name of (21).
$\Lambda(\mathbf{k}, u)$	Fourier transform of $\tilde{\psi}(\mathbf{s}, u)$.
μ	dimensionless Laplace variable ($\equiv \bar{t}u$).
ξ	$\equiv 1/v$.
ρ	dimensionless distance ($\equiv s/s_o$).
$\bar{\sigma}(t)$	standard deviation of plume displacement.
τ	dimensionless time ($\equiv t/\bar{t}$).
$\phi(\mathbf{s}, t)$	kernel function of GME.
$\tilde{\phi}(\mathbf{s}, u)$	Laplace transform of $\phi(\mathbf{s}, t)$.
$\psi(\mathbf{s}, t)$	joint probability density function for particle transitions.
$\tilde{\psi}(\mathbf{s}, u)$	Laplace transform of $\psi(\mathbf{s}, t)$.
$\psi(t)$	$\equiv \sum_{\mathbf{s}}\psi(\mathbf{s}, t)$.
$\psi_{\mathbf{s}',\mathbf{s}}(t)$	joint probability density function for particle transitions for a single realization.
tilde	Laplace transformed quantity.
angle brackets	volume average.
double brackets	ensemble average.

ADE	advection-dispersion equation.
BC	boundary condition.
BTC	breakthrough curve.
CTRW	continuous time random walk.
FPME	Fokker-Planck with memory equation.
GME	generalized master equation.
ME	master equation.
pde	partial differential equation.
pdf	probability density function.
REV	representative elementary volume.
RFN	random fracture network.

[253] **ACKNOWLEDGMENTS.** The authors thank Shlomo Neuman and two anonymous reviewers for useful comments. The financial support of BP International (B.B. and H.S.), the Israel Science Foundation (grant 598/04) (B.B.), the European Commission-Marie Curie Individual Fellowship programme (A.C.), the program "Ramon y Cajal" of the Spanish Ministry of Science and Technology (M.D.), the Spanish ENRESA (Empresa Nacional de Residuos Radioactivos) (M.D.), CICYT project MODEST (CGL-2005-05171), and the European Union IP FUNMIG (contract 516514) (M.D.) is gratefully acknowledged. B.B. holds The Sam Zuckerberg Professorial Chair.

[254] The Editor responsible for this paper was Daniel Tartakovsky. He thanks technical reviewer Shlomo P. Neuman, an anonymous technical reviewer, and an anonymous cross-disciplinary reviewer.

REFERENCES

Abramowitz, M., and I. Stegun (1970), *Handbook of Mathematical Functions*, Dover, Mineola, N. Y.

Adams, E. E., and L. W. Gelhar (1992), Field study of dispersion in a heterogeneous aquifer: 2. Spatial moment analysis, *Water Resour. Res.*, *28*(12), 3293–3308.

Aronofsky, J. S., and J. P. Heller (1957), A diffusion model to explain mixing of flowing miscible fluids in porous media, *Trans. Am. Inst. Min. Metall. Pet. Eng.*, *210*, 345–349.

Attinger, S., M. Dentz, H. Kinzelbach, and W. Kinzelbach (1999), Temporal behavior of a solute cloud in a chemically heterogeneous porous medium, *J. Fluid Mech.*, *386*, 77–104.

Attinger, S., M. Dentz, and W. Kinzelbach (2004), Exact transverse macro dispersion coefficients for transport in heterogeneous porous media, *Stochastic Environ. Res. Risk Assess.*, *18*(1), 9–15.

Auriault, J. L., and P. M. Adler (1995), Taylor dispersion in porous media—Analysis by multiple scale expansions, *Adv. Water Resour.*, *18*, 217–226.

Baeumer, B., D. A. Benson, M. M. Meerschaert, and S. W. Wheatcraft (2001), Subordinated advection-dispersion equation for contaminant transport, *Water Resour. Res.*, *37*(6), 1543–1550.

Batchelor, G. K. (1949), Diffusion in a field of homogeneous turbulence. I. Eulerian analysis, *Austral. J. Sci. Res.*, *2*, 437–450.

Bear, J. (1972), *Dynamics of Fluids in Porous Media*, Elsevier, New York.

Benson, D. A., S. W. Wheatcraft, and M. M. Meerschaert (2000), The fractional-order governing equation of Lévy motion, *Water Resour. Res.*, *36*(6), 1413–1423.

Berkowitz, B., and Y. Bachmat (1987), A scale-dependent equation for solute transport in porous media, *Transp. Porous Media*, *3*, 199–205.

Berkowitz, B., and H. Scher (1997), Anomalous transport in random fracture networks, *Phys. Rev. Lett.*, *79*(20), 4038–4041.

Berkowitz, B., and H. Scher (1998), Theory of anomalous chemical transport in fracture networks, *Phys. Rev. E*, *57*(5), 5858–5869.

Berkowitz, B., H. Scher, and S. E. Silliman (2000), Anomalous transport in laboratory-scale, heterogeneous porous media, *Water Resour. Res.*, *36*(1), 149–158.

Berkowitz, B., J. Klafter, R. Metzler, and H. Scher (2002), Physical pictures of transport in heterogeneous media: Advection-dispersion, random walk and fractional derivative formulations, *Water Resour. Res.*, *38*(10), 1191, doi:10.1029/2001WR001030.

Bhattacharya, R., and V. K. Gupta (1990), *Dynamics of Fluids in Hierarchical Porous Media*, chap. IV, pp. 61–95, Elsevier, New York.

Bijeljic, B., and M. J. Blunt (2006), Pore-scale modeling and continuous time random walk analysis of dispersion in porous media, *Water Resour. Res.*, *42*, W01202, doi:10.1029/2005WR004578.

Bijeljic, B., A. H. Muggeridge, and M. J. Blunt (2004), Pore-scale modeling of longitudinal dispersion, *Water Resour. Res.*, *40*, W11501, doi:10.1029/2004WR003567.

Bos, F. C., and D. M. Burland (1987), Hole transport in polyvinylcarbazole—The vital importance of excitation-light intensity, *Phys. Rev. Lett.*, *58*(2), 152–155.

Bouchaud, J. P., and A. Georges (1990), Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications, *Phys. Rep.*, *195*(4–5), 127–293.

Bourgeat, A., M. Quintard, and S. Whitaker (1988), Éléments de comparaison entre la méthode d'homogénéisation et la méthode de prise de moyenne avec fermeture, *C. R. Acad. Sci., Ser. II*, *306*, 463–466.

Brenner, H. (1980), Dispersion resulting from flow through spatially periodic porous media, *Proc. R. Soc. London, Ser. A*, *297*(1430), 81–133.

Bromly, M., and C. Hinz (2004), Non-Fickian transport in homogeneous unsaturated repacked sand, *Water Resour. Res.*, *40*, W07402, doi:10.1029/2003WR002579.

Carrera, J., X. Sánchez-Vila, I. Benet, A. Medina, G. Galarza, and J. Guimerà (1998), On matrix diffusion: Formulations, solution methods, and qualitative effects, *Hydrogeol. J.*, *6*, 178–190.

Christakos, G. (2004), A sociological approach to the state of stochastic hydrogeology, *Stochastic Environ. Res. Risk Assess.*, *18*(4), 274–277.

Clincy, M., and H. Kinzelbach (2001), Stratified disordered media: Exact solutions for transport parameters and their self-averaging properties, *J. Phys. A Math. Gen.*, *34*, 7141–7152.

Cortis, A., and B. Berkowitz (2004), Anomalous transport in "classical" soil and sand columns, *Soil Sci. Soc. Am. J.*, *68*, 1539–1548.

Cortis, A., and B. Berkowitz (2005), Computing 'anomalous' contaminant transport in porous media: The CTRW MATLAB toolbox, *Ground Water*, *43*(6), 947–950.

Cortis, A., Y. Chen, H. Scher, and B. Berkowitz (2004a), Quantitative characterization of pore-scale disorder effects on transport in "homogeneous" granular media, *Phys. Rev. E*, *70*(10), 041108, doi:10.1103/PhysRevE.70.041108.

Cortis, A., C. Gallo, H. Scher, and B. Berkowitz (2004b), Numerical simulation of non-Fickian transport in geological formations with multiple-scale heterogeneities, *Water Resour. Res.*, *40*, W04209, doi:10.1029/2003WR002750.

Cunningham, J. A., C. J. Werth, M. Reinhard, and P. V. Roberts (1997), Effects of grain-scale mass transfer on the transport of volatile organics through sediments: 1. Model development, *Water Resour. Res.*, *33*(12), 2713–2726.

Cushman, J. (1984), On unifying the concepts of scale, instrumentation, and stochastics in the development of multiphase transport theory, *Water Resour. Res.*, *20*(11), 1668–1676.

Cushman, J. (1991), On diffusion in fractal media, *Water Resour. Res.*, *27*(4), 643–644.

Cushman, J., and T. Ginn (1993), Non-local dispersion in media with continuously evolving scales, *Transp. Porous Media*, *13*, 123–138.

Cushman, J. H., and T. R. Ginn (2000), Fractional advective dispersion equations: A classical mass balance with convolution-Fickian flux, *Water Resour. Res.*, *36*, 3763–3766.

Cushman, J. H., and M. Moroni (2001), Statistical mechanics with three-dimensional particle tracking velocimetry experiments in

the study of anomalous dispersion. I. Theory, *Phys. Fluids*, *13*, 75–80.

Cushman, J. H., X. Hu, and T. R. Ginn (1994), Nonequilibrium statistical mechanics of preasymptotic dispersion, *J. Stat. Phys.*, *75*(5/6), 859–878.

Cushman, J. H., B. X. Hu, and F.-W. Deng (1995), Nonlocal reactive transport with physical and chemical heterogeneity: Localization errors, *Water Resour. Res.*, *31*(9), 2219–2237.

Cushman, J. H., L. S. Bennethum, and B. X. Hu (2002), A primer on up-scaling tools for porous media, *Adv. Water Resour.*, *25*, 1043–1067.

Dagan, G. (1984), Solute transport in heterogenous porous formations, *J. Fluid Mech.*, *145*, 151–177.

Dagan, G. (1987), Theory of solute transport by groundwater, *Annu. Rev. Fluid Mech.*, *19*, 183–215.

Dagan, G. (1988), Time-dependent macrodispersion for solute transport in anisotropic heterogeneous aquifers, *Water Resour. Res.*, *24*(9), 1491–1500.

Dagan, G. (1989), *Flow and Transport in Porous Formations*, Springer, New York.

Dagan, G. (1990), Transport in heterogeneous porous formations: Spatial moments, ergodicity, and effective dispersion, *Water Resour. Res.*, *26*(10), 1287–1290.

Dagan, G. (1991), Dispersion of a passive scalar in non-ergodic transport, *J. Fluid Mech.*, *233*, 197–210.

Dagan, G. (1994a), The significance of heterogeneity of evolving scales to transport in porous formations, *Water Resour. Res.*, *30*(12), 3327–3336.

Dagan, G. (1994b), An exact nonlinear correction to transverse macrodispersivity for transport in heterogenous formations, *Water Resour. Res.*, *30*(10), 2699–2705.

Dagan, G. (2004), On application of stochastic modeling of groundwater flow and transport, *Stochastic Environ. Res. Risk Assess.*, *18*(4), 266–267.

Dagan, G., and E. Bresler (1979), Solute dispersion in unsaturated soil at field scale, I, Theory, *Soil Sci. Soc. Am. J.*, *43*, 461–466.

Dagan, G., and S. P. Neuman (Eds.) (1997), *Subsurface Flow and Transport: A Stochastic Approach*, Cambridge Univ. Press, New York.

de Hoog, F. R., J. H. Knight, and A. N. Stokes (1982), An improved method for numerical inversion of Laplace transforms, *SIAM J. Sci. Stat. Comput.*, *3*, 357–366.

Deng, F.-W., J. H. Cushman, and J. W. Delleur (1993), A fast Fourier transform stochastic analysis of the contaminant transport problem, *Water Resour. Res.*, *29*(9), 3241–3247.

Dentz, M., and B. Berkowitz (2003), Transport behavior of a passive solute in continuous time random walks and multirate mass transfer, *Water Resour. Res.*, *39*(5), 1111, doi:10.1029/2001WR001163.

Dentz, M., and B. Berkowitz (2005), Exact effective transport dynamics in a one-dimensional random environment, *Phys. Rev. E*, *72*(3), 031110, doi:10.1103/PhysRevE.72.031110.

Dentz, M., and J. Carrera (2003), Effective dispersion in temporally fluctuating flow through a heterogeneous medium, *Phys. Rev. E*, *68*(3), 036310, doi:10.1103/PhysRevE.68.036310.

Dentz, M., H. Kinzelbach, S. Attinger, and W. Kinzelbach (2000), Temporal behavior of a solute cloud in a heterogeneous porous medium: 1. Point-like injection, *Water Resour. Res.*, *36*(12), 3591–3604.

Dentz, M., H. Kinzelbach, S. Attinger, and W. Kinzelbach (2003), Numerical studies of the transport behavior of a passive solute in a two-dimensional incompressible random flow field, *Phys. Rev. E*, *67*(4), 046306, doi:10.1103/PhysRevE.67.046306.

Dentz, M., A. Cortis, H. Scher, and B. Berkowitz (2004), Time behavior of solute transport in heterogeneous media: Transition from anomalous to normal transport, *Adv. Water Resour.*, *27*, 155–173, doi:10.1016/j.advwatres.2003.11.002.

Di Donato, G., E.-O. Obi, and M. J. Blunt (2003), Anomalous transport in heterogeneous media demonstrated by streamline-based simulation, *Geophys. Res. Lett.*, *30*(12), 1608, doi:10.1029/2003GL017196.

Di Federico, V., and Y.-K. Zhang (1999), Solute transport in heterogeneous porous media with long-range correlations, *Water Resour. Res.*, *35*(10), 3185–3191.

Dorfman, K. D., and H. Brenner (2002), Generalized Taylor-Aris dispersion in discrete spatially periodic networks: Microfluidic applications, *Phys. Rev. E*, *65*(2), 021103, doi:10.1103/PhysRevE.65.021103.

Eberhard, J. (2004), Approximations for transport parameters and self-averaging properties for point-like injections in heterogeneous media, *J. Phys. A Math. Gen.*, *37*, 2549–2571.

Eggleston, J., and S. Rojstaczer (1998), Identification of large-scale hydraulic conductivity trends and the influence of trends on contaminant transport, *Water Resour. Res.*, *34*(9), 2155–2168.

Einstein, A. (1905), Über die von der molekulartheoretischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen, *Ann. Phys. Leipzig*, *17*, 549–560.

Feehley, C. E., C. Zheng, and F. J. Molz (2000), A dual-domain mass transfer approach for modeling solute transport in heterogeneous aquifers: Application to the macrodispersion experiment (MADE) site, *Water Resour. Res.*, *36*(9), 2501–2515.

Freeze, R. A. (2004), The role of stochastic hydrogeological modeling in real-world engineering applications, *Stochastic Environ. Res. Risk Assess.*, *18*(4), 286–289.

Gelhar, L. W. (1993), *Stochastic Subsurface Hydrology*, Prentice-Hall, Upper Saddle River, N. J.

Gelhar, L. W. (1997), Perspectives on field-scale application of stochastic subsurface hydrology, in *Subsurface Flow and Transport: A Stochastic Approach*, edited by G. Dagan and S. P. Neuman, pp. 157–176, Cambridge Univ. Press, New York.

Gelhar, L. W., and C. L. Axness (1983), Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resour. Res.*, *19*(1), 161–180.

Gelhar, L. W., C. Welty, and K. R. Rehfeldt (1992), A critical review of data on field-scale dispersion in aquifers, *Water Resour. Res.*, *28*(7), 1955–1974.

Ghodrati, M., and W. A. Jury (1992), A field study of the effects of water application method and surface preparation method on preferential flow of pesticides in unsaturated soil, *J. Contam. Hydrol.*, *11*, 101–125.

Ginn, T. R. (2004), On the application of stochastic approaches in hydrogeology, *Stochastic Environ. Res. Risk Assess.*, *18*(4), 282–284.

Glueckauf, E. (1955), Theory of chromatography, part 10—Formulae for diffusion into spheres and their application to chromatography, *Trans. Faraday Soc.*, *51*, 1540–1551.

Guadagnini, A., and S. Neuman (2001), Recursive conditional moment equations for advective transport in randomly heterogeneous velocity fields, *Transp. Porous Media*, *42*, 37–67.

Haggerty, R., and S. M. Gorelick (1995), Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity, *Water Resour. Res.*, *31*(10), 2383–2400.

Haggerty, R., S. A. McKenna, and L. C. Meigs (2000), On the late time behavior of tracer test breakthrough curves, *Water Resour. Res.*, *36*(12), 3467–3479.

Harvey, C. F., and S. M. Gorelick (1995), Temporal moment-generating equations: Modeling transport and mass transfer in heterogeneous aquifers, *Water Resour. Res.*, *31*(8), 1895–1911.

Hatano, Y., and N. Hatano (1998), Dispersive transport of ions in column experiments: An explanation of long-tailed profiles, *Water Resour. Res.*, *34*(5), 1027–1033.

Hoffman, F., D. Ronen, and Z. Pearl (1996), Evaluation of flow characteristics of a sand column using magnetic resonance imaging, *J. Contam. Hydrol.*, *22*, 95–107.

Hornung, G., B. Berkowitz, and N. Barkai (2005), Morphogen gradient formation in a complex environment: An anomalous diffusion model, *Phys. Rev. E*, *72*(4), 041916, doi:10.1103/PhysRevE.72.041916.

Hu, B. X., F.-W. Deng, and J. H. Cushman (1995), Nonlocal reactive transport with physical and chemical heterogeneity: Linear nonequilibrium sorption with random k_d, *Water Resour. Res.*, *31*(9), 2239–2252.

Jaekel, U., and H. Vereecken (1997), Renormalization group analysis of macrodispersion in a directed random flow, *Water Resour. Res.*, *33*(10), 2287–2299.

Jiménez-Hornero, F. J., J. V. Giráldez, A. Laguna, and Y. Pachepsky (2005), Continuous time random walks for analyzing the transport of a passive tracer in a single fissure, *Water Resour. Res.*, *41*, W04009, doi:10.1029/2004WR003852.

Jury, W. A. (1982), Simulation of solute transport using a transfer function model, *Water Resour. Res.*, *18*(2), 363–368.

Jury, W. A., G. Sposito, and R. E. White (1986), A transfer function model of solute movement through soil: 1. Fundamental concepts, *Water Resour. Res.*, *22*(2), 243–247.

Kenkre, V. M., E. W. Montroll, and M. F. Shlesinger (1973), Generalized master equations for continuous-time random walks, *J. Stat. Phys.*, *9*(1), 45–50.

Kirchner, J. W., X. Feng, and C. Neal (2000), Fractal stream chemistry and its implications for contaminant transport in catchments, *Nature*, *403*, 524–527.

Kitanidis, P. K. (1988), Prediction by the method of moments of transport in heterogeneous formations, *J. Hydrol.*, *102*, 453–473.

Klafter, J., and R. Silbey (1980a), On electronic energy transfer in disordered systems, *J. Phys. Chem.*, *72*(2), 843–848.

Klafter, J., and R. Silbey (1980b), Derivation of continuous-time random walk equations, *Phys. Rev. Lett.*, *44*(2), 55–58.

Klafter, J., A. Blumen, and M. F. Shlesinger (1987), Stochastic pathway to anomalous diffusion, *Phys. Rev. A*, *35*(7), 3081–3085.

Koch, D. L., and J. F. Brady (1987), A non-local description of advection-diffusion with application to dispersion in porous media, *J. Fluid Mech.*, *180*, 387–403.

Koch, D. L., and J. F. Brady (1988), Anomalous diffusion in heterogeneous porous media, *Phys. Fluids A*, *31*, 965–1031.

Koch, D. L., and E. S. G. Shaqfeh (1992), Averaged equation and diagrammatic approximations to the average concentration of a tracer dispersed by a Gaussian random velocity field, *Phys. Fluids A*, *4*, 887–894.

Kosakowski, G. (2004), Anomalous transport of colloids and solutes in a shear zone, *J. Contam. Hydrol.*, *72*, 23–46, doi:10.1016/j.jconhyd.2003.10.005.

Kosakowski, G., B. Berkowitz, and H. Scher (2001), Analysis of field observations of tracer transport in a fractured till, *J. Contam. Hydrol.*, *47*, 29–51.

Kraichnan, R. H. (1959), The structure of isotropic turbulence at very high Reynolds numbers, *J. Fluid Mech.*, *5*, 497–543.

Kreft, A., and A. Zuber (1978), On the physical meaning of the dispersion equation and its solutions for different initial and boundary conditions, *Chem. Eng. Sci.*, *33*, 1471–1480.

Krylov, V. I., and N. S. Skoblya (1977), *A Handbook of Methods of Approximate Fourier Transformation and Inversion of the Laplace Transform*, MIR, Moscow.

Labolle, E. M., and G. E. Fogg (2001), Role of molecular diffusion in contaminant migration and recovery in an alluvial aquifer system, *Transp. Porous Media*, *42*, 155–179.

Lallemand-Barres, P., and P. Peaudecerf (1978), Recherche des relations entre la valeur de la dispersivité macroscopique d'un aquifère, ses autres caractéristiques et les conditions de mesure, *Bull. Bur. Rech. Geol. Min. Fr., Sect. 3*, *4*, 277–284.

Levy, M., and B. Berkowitz (2003), Measurement and analysis of non-Fickian dispersion in heterogeneous porous media, *J. Contam. Hydrol.*, *64*, 203–226.

Lowe, C. P., and D. Frenkel (1996), Do hydrodynamic dispersion coefficients exist?, *Phys. Rev. Lett.*, *77*(22), 4552–4555.

Lu, S. L., F. J. Molz, G. E. Fogg, and J. W. Castle (2002), Combining stochastic facies and fractal models for representing natural heterogeneity, *Hydrogeol. J.*, *10*, 475–482.

Lunati, I., S. Attinger, and W. Kinzelbach (2002), Macrodispersivity for transport in arbitrary nonuniform flow fields: Asymptotic and preasymptotic results, *Water Resour. Res.*, *38*(10), 1187, doi:10.1029/2001WR001203.

Margolin, G., and B. Berkowitz (2002), Spatial behavior of anomalous transport, *Phys. Rev. E*, *65*(3), 031101, doi:10.1103/PhysRevE.65.031101.

Margolin, G., M. Dentz, and B. Berkowitz (2003), Continuous time random walk and multirate mass transfer modeling of sorption, *Chem. Phys.*, *295*, 71–80.

Matheron, M., and G. de Marsily (1980), Is transport in porous media always diffusive?, *Water Resour. Res.*, *16*, 901–917.

Metzler, R. (2000), Generalized Chapman-Kolmogorov equation: A unifying approach to the description of anomalous transport in external fields, *Phys. Rev. E*, *62*(5), 6233–6245.

Metzler, R., and J. Klafter (2000), The random walk's guide to anomalous diffusion: A fractional dynamics approach, *Phys. Rep.*, *339*(1), 1–77.

Metzler, R., and J. Klafter (2004), The restaurant at the end of the random walk: Recent developments in fractional dynamics of anomalous transport processes, *J. Phys. A*, *37*, R161–R208.

Metzler, R., J. Klafter, and I. M. Sokolov (1998), Anomalous transport in external fields: Continuous time random walks and fractional diffusion equations extended, *Phys. Rev. E*, *58*(2), 1621–1633.

Molz, F. (2004), A rational role for stochastic concepts in subsurface hydrology: A personal perspective, *Stochastic Environ. Res. Risk Assess.*, *18*(4), 278–279.

Montroll, E. W., and H. Scher (1973), Random walks on lattices. IV. Continuous time random walks and influence of absorbing boundaries, *J. Stat. Phys.*, *9*(2), 101–135.

Montroll, E. W., and G. H. Weiss (1965), Random walks on lattices. II, *J. Math. Phys.*, *6*(2), 167–181.

Mori, H. (1965), Transport collective motion and Brownian motion, *Prog. Theor. Phys.*, *33*(3), 423–455.

Naff, R. L. (1990), On the nature of the dispersive flux in saturated heterogenous porous media, *Water Resour. Res.*, *26*(5), 1013–1026.

National Research Council (1996), *Rock Fractures and Fluid Flow: Contemporary Understanding and Applications*, Natl. Acad. Press, Washington D. C.

Neretnieks, I. (1980), Diffusion in the rock matrix: An important factor in radionuclide transport?, *J. Geophys. Res.*, *85*(B8), 4379–4397.

Neuman, S., and Y. Zhang (1990), A quasi-linear theory of non-Fickian and Fickian subsurface dispersion: 1. Theoretical analysis with application to isotropic media, *Water Resour. Res.*, *26*(5), 887–902.

Neuman, S. P. (1993), Eulerian-Lagrangian theory of transport in space-time nonstationary velocity fields: Exact nonlocal formalism by conditional moments and weak approximation, *Water Resour. Res.*, *29*(3), 633–645.

Neuman, S. P. (2004), Stochastic groundwater models in practice, *Stochastic Environ. Res. Risk Assess.*, *18*(4), 268–270.

Neuman, S. P., C. L. Winter, and C. M. Newman (1987), Stochastic theory of field-scale Fickian dispersion in anisotropic porous media, *Water Resour. Res.*, *23*(3), 453–466.

Nielsen, D. R., and J. W. Biggar (1962), Miscible displacement in soils: III. Theoretical considerations, *Soil Sci. Soc. Am. Proc.*, *26*, 216–221.

Nkedi-Kizza, P., J. W. Biggar, H. M. Selim, M. T. van Genuchten, P. J. Wierenga, J. M. Davidson, and D. R. Nielsen (1984), On the equivalence of two conceptual models for describing ion exchange during transport through an aggregated oxisol, *Water Resour. Res.*, *20*(8), 1123–1130.

Noolandi, J. (1977), Multiple-trapping model of anomalous transit-time dispersion in *a*-Se, *Phys. Rev. B*, *16*, 4466–4473.

O'Brien, G. S., C. J. Bean, and F. McDermott (2003a), Numerical investigations of passive and reactive flow through generic single fractures with heterogeneous permeability, *Earth Planet. Sci. Lett.*, *213*(3–4), 271–284.

O'Brien, G. S., C. J. Bean, and F. McDermott (2003b), A numerical study of passive transport through fault zones, *Earth Planet. Sci. Lett.*, *214*(3–4), 633–643.

Oppenheim, I., K. E. Shuler, and G. H. Weiss (1977), *Stochastic Processes in Chemical Phsyics: The Master Equation*, MIT Press, Cambridge, Mass.

Oswald, S., W. Kinzelbach, A. Greiner, and G. Brix (1997), Observation of flow and transport processes in artificial porous media via magnetic resonance imaging in three dimensions, *Geoderma*, *80*, 417–429.

Pfister, G., and H. Scher (1978), Non-Gaussian transient transport in disordered solids, *Adv. Phys.*, *27*, 747–798.

Quintard, M., and S. Whitaker (1994), Transport in ordered and disordered porous media I: The cellular average and the use of weighting functions, *Transp. Porous Media*, *14*, 163–177.

Rajaram, H., and L. W. Gelhar (1993), Plume-scale dependent dispersion in heterogeneous aquifers: 2. Eulerian analysis and three-dimensional aquifers, *Water Resour. Res.*, *29*(9), 3261–3276.

Rajaram, H., and L. W. Gelhar (1995), Plume-scale dependent dispersion in aquifers with a wide range of scales of heterogeneity, *Water Resour. Res.*, *31*(10), 2469–2482.

Rao, P. S. C., R. E. Jessup, D. E. Rolston, J. M. Davidson, and D. P. Kilcrease (1980), Experimental and mathematical description of nonadsorbed solute transfer by diffusion in spherical aggregates, *Soil Sci. Soc. Am. J.*, *44*(4), 684–688.

Roberts, P. H. (1961), Analytical theory of turbulent diffusion, *J. Fluid Mech.*, *11*, 257–283.

Roth, K., and W. A. Jury (1993), Linear transport models for adsorbing solutes, *Water Resour. Res.*, *29*(4), 1195–1203.

Rubin, Y. (2003), *Applied Stochastic Hydrogeology*, Oxford Univ. Press, New York.

Rubin, Y. (2004), Stochastic hydrogeology: Challenges and misconceptions, *Stochastic Environ. Res. Risk Assess.*, *18*(4), 280–281.

Rubinstein, J., and R. Mauri (1983), Dispersion and convection in periodic porous media, *SIAM J. Appl. Math.*, *46*(6), 1018–1023.

Sardin, M., D. Schweich, F. J. Leij, and M. T. van Genuchten (1991), Modeling the nonequilibrium transport of linearly interacting solutes in porous media: A review, *Water Resour. Res.*, *27*(9), 2287–2307.

Scheidegger, A. E. (1959), An evaluation of the accuracy of the diffusivity equation for describing miscible displacement in porous media, *Proc. Theory Fluid Flow Porous Media Conf., 2nd*, 101–116.

Scher, H., and M. Lax (1973a), Stochastic transport in a disordered solid. I. Theory, *Phys. Rev. B*, *7*, 4491–4502.

Scher, H., and M. Lax (1973b), Stochastic transport in a disordered solid. II. Impurity conduction, *Phys. Rev. B*, *7*, 4502–4519.

Scher, H., and E. W. Montroll (1975), Anomalous transit time dispersion in amorphous solids, *Phys. Rev. B*, *12*, 2455–2477.

Scher, H., G. Margolin, and B. Berkowitz (2002a), Towards a unified framework for anomalous transport in heterogeneous media, *Chem. Phys.*, *284*, 349–359.

Scher, H., G. Margolin, R. Metzler, J. Klafter, and B. Berkowitz (2002b), The dynamical foundation of fractal stream chemistry: The origin of extremely long retention times, *Geophys. Res. Lett.*, *29*(5), 1061, doi:10.1029/2001GL014123.

Schmidlin, F. W. (1977), Theory of trap-controlled transient photoconduction, *Phys. Rev. B*, *6*, 2362–2385.

Schumer, R., D. A. Benson, and M. M. Meerschaert (2003), Fractal mobile/immobile solute transport, *Water Resour. Res.*, *39*(10), 1296, doi:10.1029/2003WR002141.

Shlesinger, M. F. (1974), Asymptotic solutions of continuous-time random walks, *J. Stat. Phys.*, *10*(5), 421–434.

Shlesinger, M. F. (1988), Fractal time in condensed matter, *Annu. Rev. Phys. Chem.*, *39*, 269–290.

Shlesinger, M. F. (1996), *Random Processes, Encyclopedia of Applied Physics*, vol. 16, John Wiley, Hoboken, N. J.

Shlesinger, M. F., G. M. Zaslavsky, and J. Klafter (1993), Strange kinetics, *Nature*, *363*, 31–37.

Shvidler, M. J. (1993), Correlation model of transport in random fields, *Water Resour. Res.*, *29*, 3189–3199.

Sidle, C., B. Nilson, M. Hansen, and J. Fredericia (1998), Spatially varying hydraulic and solute transport characteristics of a fractured till determined by field tracer tests, Funen, Denmark, *Water Resour. Res.*, *34*(10), 2515–2527.

Silliman, S. E., and E. S. Simpson (1987), Laboratory evidence of the scale effect in dispersion of solutes in porous media, *Water Resour. Res.*, *23*(8), 1667–1673.

Sposito, G., R. E. White, P. R. Darrah, and W. A. Jury (1986), A transfer function model of solute transport through soil: 3. The convection dispersion equation, *Water Resour. Res.*, *22*(2), 255–262.

Stark, C. P., and M. Stieglitz (2000), The sting in a fractal tail, *Nature*, *403*, 493–495.

Sudicky, E. (2004), On certain stochastic hydrology issues, *Stochastic Environ. Res. Risk Assess.*, *18*(4), 285.

Tiedje, T. (1984), Information about band-tail states from time-of-flight experiments, in *Semiconductors and Semimetals, Part C*, vol. 21, pp. 207–238, Elsevier, New York.

Villermaux, J. (1974), Deformation of chromatographic peaks under the influence of mass transfer phenomena, *J. Chromatogr. Sci.*, *12*, 822–831.

Villermaux, J. (1987), Chemical engineering approach to dynamic modelling of linear chromatography, *J. Chromatogr.*, *406*, 11–26.

Weisbrod, N., M. R. Niemet, and J. Selker (2003), Light transmission technique for the evaluation of colloidal transport and dynamics in porous media, *Environ. Sci. Technol.*, *37*(16), 3694–3700.

Weissmann, G. S., S. F. Carle, and G. E. Fogg (1999), Three-dimensional hydrofacies modeling based on soil surveys and transition probability geostatistics, *Water Resour. Res.*, *35*(6), 1761–1770.

Whitaker, S. (1999), *The Method of Volume Averaging*, Springer, New York.

Winter, C. L. (2004), Stochastic hydrology: Practical alternatives exist, *Stochastic Environ. Res. Risk Assess.*, *18*(4), 271–273.

Zhang, D. (2002), *Stochastic Methods for Flow in Porous Media: Coping With Uncertainties*, Elsevier, New York.

Zhang, D., and S. Neuman (1995), Eulerian-Lagrangian analysis of transport conditioned on hydraulic data: 1. Analytical-numerical approach, *Water Resour. Res.*, *31*(1), 39–52.

Zhang, Q. (1995), Transient behavior of mixing induced by a random velocity field, *Water Resour. Res.*, *31*, 2955–2963.

Zhang, X., J. W. Crawford, L. K. Deeks, M. I. Stutter, A. G. Bengough, and I. M. Young (2005), A mass balance based numerical method for the fractional advection-dispersion equation: Theory and application, *Water Resour. Res.*, *41*, W07029, doi:10.1029/2004WR003818.

Zhang, Y., and D. Zhang (1996), Nonergodic solute transport in three-dimensional heterogeneous isotropic aquifers, *Water Resour. Res.*, *32*, 2955–2963.

Zhang, Y.-K., and D. Zhang (2004), Forum: The state of stochastic hydrology, *Stochastic Environ. Res. Risk Assess.*, *18*(4), 265.

Zwanzig, R. (1960), Ensemble method in the theory of irreversibility, *J. Chem. Phys.*, *33*(5), 1338–1341.

B. Berkowitz and H. Scher, Department of Environmental Sciences and Energy Research, Weizmann Institute of Science, Rehovot 76100, Israel. (brian.berkowitz@weizmann.ac.il; harvey.scher@weizmann.ac.il)

A. Cortis, Earth Sciences Division 90-1116, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA. (acortis@lbl.gov)

M. Dentz, Department of Geotechnical Engineering and Geosciences, Technical University of Catalonia (UPC), Barcelona E-08034, Spain. (marco.dentz@upc.es)